THE UNIVERSITY *of* EDINBURGH

Competition between specialist and generalist species in computational and experimental model ecosystems

Catherine Mills

Doctor of Philosophy
The University of Edinburgh
February 2017

Declaration

I declare that this thesis has been composed solely by myself and that it has not been submitted, in whole or in part, in any previous application for a degree. Except where states otherwise by reference or acknowledgment, the work presented is entirely my own.

Acknowledgements

Thanks to my supervisors Richard Blythe and Rosalind Allen for what felt like a perfect balance of freedom and guidance.

Thanks to everyone in Edinburgh who made my time here what it was, but especially Toby for feeding me.

It is what it is.

Lay Summary

The natural world contains many species. These species interact in various ways. The study of these interactions and their effects is called ecology. Many species that coexist in the same geographical area can be thought of as an ecological community. These communities can look quite different. The reasons why some communities look different to other communities are still mostly unknown.

Species are suited to their environment to a varying degree. Species that are highly suited to a particular environment, but ill-suited to others, are called specialists; those that are moderately suited to all environments are called generalists. We are interested in why different communities have different amounts of specialist and generalist species.

We model competition between specialist and generalist species using a computer simulation with imaginary, idealised species that compete with each other for imaginary, idealised resources on a grid. Firstly we simulate using a 'well-mixed' environment where we assume all species can travel to any other place on the grid to compete with one another instantly. We find that, as the simulation is run over time, if the population of specialists in the environment becomes large enough, this causes the generalist population to suddenly crash and for the whole environment to become full of specialists. In other words, the presence of specialists allows other specialists to live in the same area.

Secondly, we change the simulation so that now species can only travel a certain distance to compete with each other, and different species can travel further than other species. We find that generalist species are usually more successful if they can travel a long way to complete, and specialist species are more successful if they travel a short way.

In addition to investigating specialists and generalists using computer simulations, we also investigate them using real-world experiments of micro-organisms, which

are very small organisms that are invisible to the naked eye.

Micro-organisms form an interdependent community in which different species perform different tasks that keep the community functioning. For example, some micro-organisms degrade waste products from other micro-organisms, returning nutrients to the system that can then be reused by other micro-organisms. We can therefore group micro-organisms according to the functions they perform in the community. After some analysis, we find that different micro-organisms have a specific set of other micro-organisms that they 'prefer' to 'work' with. We also find that generalists appear to be more dependent on one another than specialists.

Abstract

An ecological community is complex and the mechanisms behind the assembly of such a community are still poorly understood. Here, we concentrate on the question of what mechanisms affect the proportion of specialists and the proportion of generalists in a community.

First, we use an individual-based model to explore the effects of the available resource spectrum on the specialist-generalist balance in well-mixed and spatially structured environments. In the well-mixed model, we uncover a new mechanism which we term 'resource spectrum engineering', in which opportunistic specialists occupying small niches in a mostly generalist community can change the resource spectrum that is experienced by other species strongly disfavouring generalists and causing a community-wide shift towards specialist strategies. Extending to a spatially structured model in which the dispersal distance of species may be limited, we find that specialism is linked to intermediate dispersal lengths, whereas generalism is linked to short and long dispersal lengths.

We then investigate two real microbial systems, using 16rRNS sequence data. In the first experiment, we identify functional groups of specialists and generalists by perturbing the microbial environment with variable nutrient concentrations and establishing which groups survive across different concentrations and which do not. In the second experiment we use many replicates of samples from the same source to find co-occurrence correlations between species, which suggest mutualistic relationships. We find that generalist species may be more likely to be dependant on the presence of each other than on specific environmental conditions.

Contents

List of Figures

Chapter 1

Introduction

In this chapter, we will outline a number of important ecological concepts that will be important for the rest of the thesis. First, we will discuss what an ecosystem is, how a natural ecosystem may be examined, and different types of interactions between species. Next, we will introduce the important concept of a *niche*, which is the specific role that a species may play in an ecosystem. Next, we introduce the concepts of specialists and generalists, and outline the effect on balance between specialists and generalists that different types of environments may have. Finally, we discuss previous attempts at modelling specialists and generalists.

1.1 Ecosystems

An ecosystem can be defined as a number of species that exist together in the same physical area and influence each other. An ecosystem can and often does comprise of species of a huge variety of spatial scales, which are nevertheless essential to the survival of one another; for example the presence of bacteria that degrade organic matter in the gut of an elephant is crucial to the survival of the elephant, just as the survival of the elephant is essential to the survival of the bacterial community that lives inside it.

It is very often not possible to draw definitive borders around a single ecosystem. Apart from experimental model ecosystems (such as a community of bacteria in a test tube) and rare cases of very isolated ecosystems in the real world such as communities based around geothermal vents, the only true border that can be

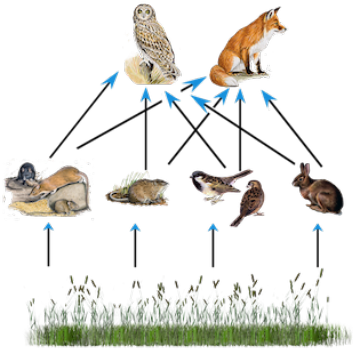


Figure 1.1 *An illustration of a food web with different trophic levels (reproduced from [1])*

used to define an entire ecosystem is that of the entire planet. Even comparatively isolated land areas such as Australia are still connected with the rest of the world by the presence of fishes in the ocean and birds in the air, which will have an effect on the land-dwelling species.

It must therefore be accepted that when studying ecology and when attempting to model ecosystems, compromises must be made. For example, although an island is not truly isolated, it can be assumed to be isolated for many intents and purposes. Similarly, although a very large number of species may be essential to the functioning of an ecosystem, it is possible to select a subset of those species that are relevant to a particular ecological question and study only them. However, one must always bear in mind that doing this is a simplification of the system that truly exists and care must be taken to minimise the effect of doing this.

1.1.1 Trophic levels

Species in an ecosystem can be divided into different trophic levels. This is a term which refers to the place in the food chain that species occupies. The bottom trophic level contains producers, being species such as plants and algae that obtain energy from the sun by photosynthesis. On the next level, primary consumers are herbivorous species that feed on the producers, which includes certain insects, rodents and larger herbivores such as cows. The secondary consumers then feed on the primary consumers, such as rodents and birds that feed on insects. Classic predators such as birds of prey and cats then feed on

them. Figure 1.1 shows a very simple illustration of an ecosystem with three trophic levels.

Detrivores are species that feed on the dead bodies of other species, such as vultures and striped hyenas. Decomposers such as fungi and degrading bacteria then decay whatever is unpalatable such that the nutrients may be returned to the ground or sea, and used again by the producers.

Clearly, these trophic levels are not always clearly defined and there may be varying numbers of species in an individual food chain. For example, a mouse may feed on both insects and grain, making it both a primary and secondary consumer, as may many other omnivores. A top predator may feed on only herbivorous prey (such as a lion eating a gazelle), or it may feed on prey that are many degrees of separation away from a producers (such as sharks, which eat tuna that have eaten mackerel that have eaten smaller fish that have eaten zooplankton that have eaten phytoplankton). Nevertheless, it is useful to group species together that all eat roughly the same thing, and that are all at risk from similar predators. For example, mice, squirrels, bats and small birds can all be considered to be on the same trophic levels as they eat roughly the same things (seeds, berries and insects) and are threatened by the same predators (birds of prey, snakes, cats). It is therefore meaningful to compare these species and the strategies they use.

A group of species on the same trophic level which use a particular type of resource in a similar way are referred to as a guild [20, 32, 143]. Species may or may not be genetically related. The theoretical models in this thesis model competition between a single guild of species.

1.1.2 Interactions between species

Understanding interspecies interactions is crucial to understanding how an ecosystem works. There are many ways in which species can interact, but these can fortunately be grouped into a small number of simple categories.

Figure 1.2 shows a table of possible interactions between two species, denoted 1 and 2.

		Species 2		
		+	0	−
Species 1	+	Mutualism	Commensalism	Antagonism
	0	Commensalism	No effect	Amensalism
	−	Antagonism	Amensalism	Competition

Figure 1.2 *Possible interactions between species*

(+,+) Mutualism

If species interact in a way that is favourable to both, the resulting interaction is termed mutualistic. For example, a mutualistic relationship exists between cleaner fish and their hosts. A cleaner fish travels with its host, eating parasites that collect on it. The cleaner fish benefits by having a steady supply of food, and by having the protection of its much larger host, while the host benefits by having its parasites conveniently removed [27]. Some pairs of species depend on one another so much that they cannot live without one another; this mutualistic relationship then becomes symbiotic.

(+,0) Commensalism

A commensal relationship is when one species benefits from the presence of another, but the other species is unaffected. For example, barnacles traveling on a whale benefit from the safety of staying with a huge whale, while the whale is unaffected by the presence of the barnacles [13].

(+,-) Antagonism

An antagonistic relationship is where one species benefits from an interaction while the other is harmed. This includes all predator-prey interactions, parasite-host interactions and herbivorism [4].

(-,0) Amensalism

If one species is harmed by an interaction, while having no effect on the other, the relationship is amensalistic. For example, small insects and ibex both feed on grass [62]. While eating the grass, the ibex may eat the insects, which is highly detrimental to the insects but has negligible effect on the ibex.

(-,-) Competition

Species whose presence has a negative effect on one another are said to be in competition for some resource such as food or space. In most cases, only species that are on the same trophic level can be in competition with one another. Even if two species consume the same resource, if they are on different trophic levels (such as insects and ibex eating the same grass), they are not considered to be in competition.

In the computational models outlined in this thesis we will be focusing almost entirely on the mechanism of competition between species on a single hypothetical trophic level. We will also show that through the mechanism of competition, different species that do not appear to interact at all may in fact be interacting mutualistically or commensally.

1.1.3 Diversity

The diversity of an ecosystem is a fundamental concept in ecology, meaning the range of different species present. Diversity can be measured in a number of different ways. Firstly, one can to simply count the species that appear in an ecosystem with a nonzero abundance. This is known as the 'species richness'. However, this can give a misleading measure of diversity if one species has a very high abundance and there are many species with low abundance. The high number of species counted in this system would result in a high species richness, but give no indication that one species is dominating.

To avoid this, throughout this thesis we have used the Shannon diversity to calculate diversity, although other diversity measures exist. This takes into account both the number of species counted and the abundance of each species, such that a large number of species with a similar abundance would result in a large measure for diversity. The following equation shows the Shannon diversity for N species, where p_i is the fractional abundance of a species i.

$$\text{Shannon Diversity} = -\sum_{i=1}^{N} p_i \ln(p_i) \tag{1.1}$$

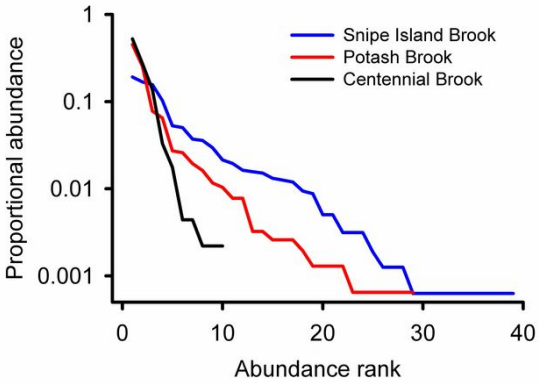


Figure 1.3 *An example of three rank abundance curves from three stream communities in Vermont (reproduced from [103])*

Rank abundance curves

A good way to visualise the diversity of an ecosystem in more detail is to plot a rank abundance curve. To plot a rank abundance curve, all species sampled in a geographical area, or ecosystem, are ranked according to their abundance (number of individuals counted). This rank is then plotted against their abundance. To provide an illustration, Figure 1.3 shows an example of three rank abundance curves from communities in three streams in Vermont. We can see that Snipe Island Brook has a higher richness than the other two streams, as it contains nearly forty species. However, we can also see that Snipe Island Brook has a more even abundance distribution than the other two streams. The species with rank 1 in Snipe Island Brook has a much lower relative abundance than those with rank 1 at Potash Brook and Centennial Brook. Snipe Island Brook is therefore more diverse in both measures. We will use rank abundance curves to visualise simulated communities in the modelling sections of this thesis.

1.1.4 Equilibrium and stability

Many ecological theories assume that an ecosystem is fundamentally in an equilibrium state, meaning that the abundance of each species is constant over time, even though it may undergo fluctuations. Even if species populations deviate a great deal from this equilibrium, as in the case of Lotka-Volterra predator-prey models, it is nonetheless often assumed that the system will continue to cycle around a single fixed point [45]. However, this is an unproven assumption. Sprugel [145] points out that equilibrium is impossible to reach in

Figure 1.4 *Two areas of a lake separated by a curtain, one side of which has undergone eutrophication (adapted from [149])*

some areas because of regular large disturbances such as forest fires, and argues that some areas where disturbances are small may reach equilibrium, but areas where disturbances are larger may not.

In this thesis, in our simulated ecosystems, we will assume that the environment is unchanging and so the community can reach an equilibrium state. Similarly, in our experimental model ecosystems, the ecosystems will on the most part be left undisturbed and permitted to reach equilibrium. However, it is important to bear in mind that the assumption of an ecosystem at equilibrium is an unproven assumption when considering real ecosystems in the natural world.

Stability

Even if an ecosystem is in an equilibrium state, it may still be possible to perturb it enough such that it transitions to an entirely different equilibrium state. A perturbation such as a destructive environmental event or a collapse in population of a key species may result in the entire ecosystem ceasing to function in the same way. An ecosystem that contains generalists that may take on each others roles interchangeably may be more stable than one that does not. For example, in a system in which two different predators consume the same prey, if one predator is eliminated through disease, the other predator can continue to keep prey numbers down, which will avoid the possible disruptive effects of an excess of prey.

Alternative ecosystem states

Ecosystems may sometimes have differing stable states, such that differing initial conditions, or a large perturbation from one state, can result in different long-time

steady states. For example, if the water of a transparent and ecologically diverse lake becomes too nutrient-rich, the lake may undergo sudden eutrophication, in which the water becomes cloudy and many plant and animal species die. Once the lake is in this state, it can be reversed back to its previous transparent state by temporary removal of fish biomass [77, 104, 136, 137]. Figure 1.4 shows the results of a famous experiment in which a curtain was placed across a lake in Canada, separating one side from the other. Different quantities of carbon, nitrogen and phosphorous were added to each side. The side with additional phosphorus underwent eutrophication while the other side remained clear, providing an example of how an ecosystem can tip into another state with the right conditions.

1.2 The niche hypothesis

The concept of a 'niche' refers to the idea that a species utilises a specific set of resources in an environment to survive. This concept plays a central role in ecological theory. A niche axis refers to a type of resource or habitat varying along a single dimension - for example, how high in a tree a species lives, or the size of food it consumes.

1.2.1 Niche partitioning

Niche partitioning refers to the phenomenon whereby similar species competing along the same niche axis differentiate themselves by occupying different parts of the axis, such that they are not in competition with one another. A famous example of niche partitioning is found in the Galápagos finches (or Darwin's finches). These are a group of closely related bird species which all occupy the same geographical area, whose beaks have evolved to be different sizes, so that each bird species can eat seeds of a different size [166].

There are many different resource 'axes' along which a species can occupy a specific niche; for example, the nutrient content of the food it consumes, the size of the food, the elevation at which it lives and the time of day at which it is awake are all examples of resource axes. The fundamental niche of an organism is therefore an 'n-dimensional hypervolume' in the space of all the different resource axes experienced by the organism [36, 72, 84, 126]. Figure 1.5 shows an illustration of six hypothetical species coexisting over two niche axes. Although two species

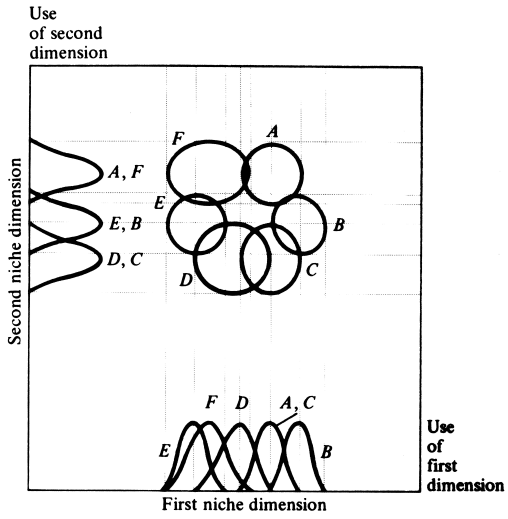


Figure 1.5 *The 2-dimensional hypervolume of a number of coexisting species (reproduced from [2])*

may occupy the same niche when considering only the first or second axes, when considered together, each species occupies its own unique niche space.

Schoener [140] provides an elegant example of niche partitioning over multiple dimensions in a study of tree dwelling lizards. As this study explains, intuitively large lizards should eat large insects, and perch on areas of the tree with thick branches, while small lizards should eat small insects, and perch on parts of the tree with thinner branches. However, this arrangement would result in lizards of similar body size inhabiting the same part of a tree, competing for the same insects. Instead, counter-intuitively, the lizards have evolved to avoid competing with one another by having the smaller species occupy the largest perches and the larger species occupy the smallest perches. As there is a direct correlation between body and perch size within a species, and an inverse correlation between species, the result is to place the largest members of a large species adjacent to the smallest members of a smaller species along the branch, where they may coexist eating different insects without competing with one another.

In our computational models we will assume that there is only a single environmental resource axis. However, it is important to understand that this is a simplifying assumption.

Regularly spaced niche partitioning

A theory in ecology states that species will partition themselves in a regular pattern along a niche axis, provided the environment is stable enough. The theory assumes that, given enough time, the maximum number of species that can coexist along a niche axis will arise. If the niche axis is maximally filled, the spacing along the axis will be regular, as species that coexist adjacent to one another along the niche axis will be as closely related as it is possible to be without excluding one another.

Some evidence exists for this theory. Studying a number of coexisting mammal and bird species, Hutchinson [73] found that the size of the skull (in the case of mammals) or culmen[1] (in the case of birds) of competing species that co-occur in the same place took a ratio varying from 1.1 to 1.4. Subsequent studies often, but not always, showed similar patterns regarding body size [140]. It is important to note that this theory is not easily falsified; a study that does not show regular niche partitioning along the investigated niche axis cannot be taken as evidence that regular partitioning is not occurring, as it is always possible that partitioning is taking place along another niche which is not being measured.

Some theoretical models have shown regularly spaced niche partitioning. Using a Lotka-Volterra competition model (the details of which will be explained in Section 1.5.3), Scheffer [138] showed that hypothetical competing species would partition themselves regularly along an arbitrary periodic niche axis. In our model, we show the same phenomenon using an individual-based model.

1.2.2 Fundamental and realised niches

Although the 'fundamental niche' of a species refers to all resources that it is possible for it to utilise, it is important to note that a species will usually not use its entire fundamental niche, but rather the 'realised niche' of what is available to it under competition [126]. The fundamental niche is the niche it occupies when under no competition constraint; the niche it occupies when a single species is placed, alone, in an environment where it is free to use whatever resources it physically can and occupy whatever space it physically can. The realised niche of a species may vary significantly depending on the ecosystem it inhabits. If

[1] The upper ridge of a bird's beak

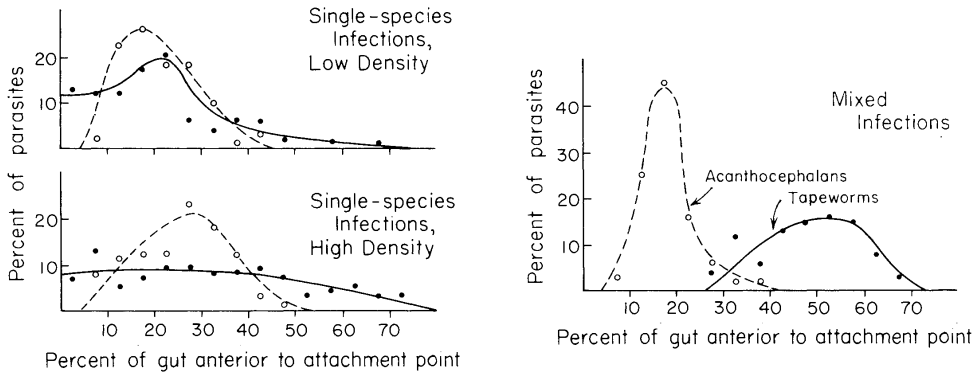


Figure 1.6 *The fundamental and realised niches of a parasitic tapeworm and acanthocephalan (adapted from [36])*

competition is removed, species will often use a larger portion of the resource gradient than they would under natural conditions.

Figure 1.6 shows an illustration of this concept using two parasites that exist in the gut. Holmes [69] performed a number of experiments with the intestinal parasites tapeworms and acanthocephalans in wild rats, to compare the regions of the intestine occupied when the parasites were alone, and when they were placed in competition with one another. The x-axis on each plot shows the position along the length of the gut where each parasite attaches. The left hand two plots show where in the gut the parasites have attached when they are alone in the gut, at low and high infection density. The right hand plot shows where the parasites attach when they are both present in the same gut. As we can see, when the parasites are alone, they occupy a much larger proportion of the gut. This represents the species' fundamental niche. When they are placed in competition with one another, the acanthocephalans occupy the anterior region of the intestine, while the tapeworms occupy the posterior region. These are the realised niches for these species in this particular situation. Many other experimental examples are outlined in a review by Colwell [36].

A measure often discussed in ecology is the 'niche breadth' of a species, which refers to the range of resources it utilises. It is important to distinguish between the fundamental niche breadth and the realised niche breadth, as they may be quite different.

In our modelling work, we will technically assume that the fundamental niche of all of our simulated species spans the entire width of the resource spectrum; if any species was placed on its own in the environment it would come to occupy the

whole environment. As species are always put in competition with one another, the realised niche is much smaller. We will find that in more specialist species, the realised niche depends on the presence of that species' ideal resource, and in more generalist species, the realised niche depends more on simply what resources are unused by other species.

1.2.3 The competitive exclusion principle

The competitive exclusion principle states that if more than one species attempts to occupy the same niche, the less fit species will be outcompeted and will become extinct [12, 65].

Thus, each species must have a wide range of axes on which to differentiate; two species may occupy the same niche along a single niche axis, but they may occupy different positions on other niche axes, resulting in different niche hypervolumes and thus not competitively excluding one another [165]. From empirical observations, a group of species competing for similar resources can be expected to differentiate on the order of three axes [140].

The Great American Interchange is a compelling story that supports the niche hypothesis and competitive exclusion. Around three million years ago, a land bridge formed between North and South America. Previous to this, the continents had been separated and entirely different ecosystems had emerged in each one. After the formation of the land bridge, many fauna and flora in South America became extinct. Many species in North and South America were occupying very similar niches, so when they were found in the same geographical location, the fitter species would outcompete the other according to the competitive exclusion principle. The species from North America were fitter (perhaps because the continent had undergone multiple invasions in the past, while South America had remained isolated for a long time), and so completely displaced those in South America [101, 124].

This principle may be derived mathematically. Following Armstrong and McGehee [11], we consider N species, competing for a single resource R, each with a growth rate γ and a death rate σ.

$$\frac{dN_i}{dt} = N_i(\gamma_i R - \sigma_i) \tag{1.2}$$

We then assume that the total resource R is diminished as it is consumed by competitors, such that available R is in fact a function of the number of competitors, subtracted from a maximum $R_{\max}$:

$$R = R_{\max} - F(N_1...N_m) \tag{1.3}$$

Substituting 1.2 into 1.3 and rearranging, we find:

$$\frac{dN_i}{dt} = N_i(\gamma_i R_{\max} - \sigma_i - \gamma_i F(N_1...N_m)) \tag{1.4}$$

We can now replace $\gamma_i R_{\max} - \sigma_i$ with ϵ_i. Rearranging results in the following:

$$\frac{dN_i}{dt} = N_i(\epsilon_i - \gamma_i F(N_1...N_m)) \tag{1.5}$$

As t approaches ∞, the species with the largest value of ϵ_i/γ_i is the only one that approaches a nonzero value. This demonstrates that if many species are competing for the same resource, with time, only the fittest will survive.

However, this derivation is subject to many simplifying assumptions, including that the growth rate of a species i is a linear function of R (Equation 1.2). Armstrong and McGehee [11] show that if this assumption is relaxed, any number of species may coexist on only 4 resources. They argue that the assumption of linearity is violated in seemingly all biological cases, and the competitive exclusion principle must therefore not be taken as proven. However, the patterns implying the competitive exclusion principle may be seen in many sets of empirical data, and it remains a useful tool as a way of thinking about patterns of species.

1.2.4 The paradox of the plankton

Given the competitive exclusion principle, why do we observe so many coexisting species in nature? There are numerous examples of many species that appear to use exactly the same resources, but nonetheless coexist in the same geographical area together. This has been observed in the case of rainforest and coral reef communities [37], but is best illustrated with the case of plankton, and particularly of phytoplankton. Hutchinson named this problem 'the paradox

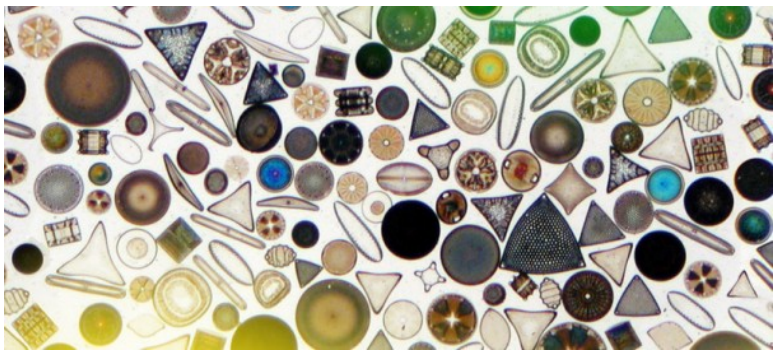


Figure 1.7 *A variety of phytoplankton (reproduced from [3])*

of the plankton' [74] in 1961. As we see in Figure 1.7, phytoplankton in the ocean is extremely diverse and includes a wide variety of different body shapes. However, all species of phytoplankton compete for very similar materials; they all require light, CO_2 and a number of similar vitamins and minerals. This therefore presents a problem for the competitive exclusion principle, as this would appear to be a case of many species occupying the same niche, and yet coexisting.

A number of explanations have been proposed to resolve the paradox of the plankton. Firstly, different species may occupy different niches that are not immediately obvious, and so are not in direct competition with one another at all. More recently, a study by Stomp et al [150] showed that two strains of picocyanobacteria (a type of phytoplankton) can stably coexist as one strain utilises the red part of the light spectrum while the other utilises the green part. Niche differentiation along the light spectrum may explain some of the diversity shown.

Secondly, a 'kill-the-winner' [174] (also known as Janzen-Connell [172]) mechanism may be at play. In this mechanism, an organism that becomes particularly successful finds itself a target for parasites or predators that target it specifically. This then results in the population of the successful organism decreasing and so disallowing the dominance of any one organism.

Thirdly, the environment and resources for which species are competing may in fact never be at equilibrium. This would prevent any well-adapted species having the time to become dominant and displace all others. At any particular time, the community is always made up of species which are adapted to the current conditions, and species which were adapted to the previous conditions. This is also known as the 'storage effect', in reference to previously successful species being stored over time until their ideal conditions occur again [172]. It is unclear

whether ecosystems are at equilibrium or not; this is very important in ecological theory and will be discussed in more detail later.

As our computational models include only the competition mechanism, and no predation or parasitism, the kill-the-winner explanation is not explored in this thesis. In addition, our simulated environments are static, and so the non-equilibrium explanation cannot be explored in full. We instead concentrate on niche differentiation as an explanation for diversity.

1.3 Specialists and generalists

Niche breadth can vary widely among different species. Species with a narrow niche breadth utilise a small part of the available resource spectrum and can be thought of as specialists. Species with a wide niche breadth utilise a large part of the resource spectrum and can be thought of as generalists. A central theme in ecology concerns the composition of an ecological community in terms of specialists and generalists.

All species have some degree of specialisation, or lack of it. A giant panda is clearly a specialist, as it depends exclusively on a single food source and cannot eat anything else. Its digestive system is adapted for one food, and one food only. Grizzly bears and pigs are clearly more generalist, being omnivorous and able to digest a wide range of foods. Humans are the greatest generalists of all, not only having the ability to digest a wide range of food but having the adaptability to survive on all seven continents, a feat not shared by any other non-microbial species.

Returning to the concept of fundamental and realised niches, we can see that the definition of a generalist may be quite different when considering different types of niche. Consider a species that is found to be utilising a particular part of the resource spectrum in an environment. By considering only the *realised* niche, this species may be thought of as a specialist. However, without examining the species in many other environments, it is not possible to know whether it is actually a generalist in terms of its *fundamental* niche. It may have the ability to utilise a wide range of resources, but is being outcompeted by more specialist species on most of them, subsisting only on the available resources. In our computational modelling work, we will see that fundamentally generalist species may subsist

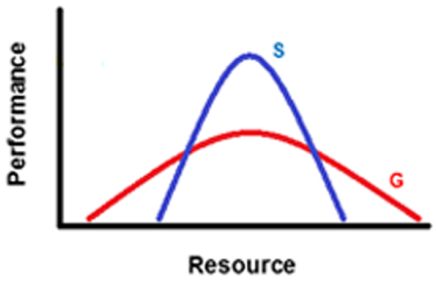


Figure 1.8 *Diagram showing the oft-assumed tradeoff between specialists (S, blue) and generalists (G, red) (adapted from [125])*

on a narrow range of resources because specialists are occupying all other parts of the resource spectrum. Additionally, in our analysis of experimental data, we will encounter microbial species that may appear to be specialist from our experiments but is actually likely to be more generalist when we also consider previous literature studies.

1.3.1 The 'jack of all trades' assumption

It is often assumed that a tradeoff exists between niche breadth and the maximum efficiency of resource utilisation: specialists can outperform generalists, but only when they are in their optimal habitats (Figure 1.8). This is referred to as the 'jack of all trades' assumption, i.e. that a jack of all trades is a master of none. This trade-off has been widely assumed by both theorists and empiricists [47, 59, 84, 95, 162, 171].

An experimental study of *E. coli* (Dykhuizen and Davies, 1980)[49] provides evidence of this trade-off. In this experiment, two strains of *E. coli*, one of which is a specialist which consumes only maltose, the other a generalist consuming both maltose and lactose, were placed in a chemostat[2] with varying proportions of maltose and lactose. Coexistence occurred at intermediate levels of maltose and lactose; if the concentration of lactose was too low, the specialist won while the generalist won if the concentration of lactose was too high. This demonstrates the trade-off because it must follow that the specialist is more adept at consuming maltose than the generalist; otherwise the generalist would displace the specialist at intermediate levels of maltose and lactose.

[2]A chemostat is a bioreactor in which fresh nutrients are constantly supplied while used up nutrients and metabolic end products are removed. This results in a steady-state microbial culture [75].

Additionally, evidence for this tradeoff can be found in a review by Fry [58] of numerous selection experiments using plant parasites. In one set of experiments, spider mites that have become optimally adapted to lima bean plants were placed on a less favourable host, cucumber plants. After several generations, the spider mites became adapted to the cucumber, and they were then returned to the lima bean. The cucumber-adapted spider mites were then less fit on the lima bean. This shows that a species cannot simply become maximally fit in every environment; to survive in in one set of conditions, it tends to become less fit in others.

However, it is important to note that this assumption, although intuitive and shown in laboratory selection experiments, may not be necessarily true in the natural environment. Some studies have found counter-examples in which some species perform well in all habitats. For example, in a study of specialist lynx and generalist bobcat by Peers et al [125], it was found that although bobcat had a wider niche (being found in a wider range of environmental conditions), the peak performance of each species was the same, which may indicate that the lynx gained no advantage from specialising on certain conditions. Additionally, in a study of specialist and generalist antlion larvae by Barkae et al [17], it was found that the generalist outperformed the specialist even in the specialist's preferred habitat, although it is possible that this specialist is simply an example of an evolutionary dead end.

In our computational models, we will assume that this trade-off holds.

1.3.2 Defining specialism and generalism

Defining a species as 'specialist' or 'generalist' when species are compared across different environments, spatial and temporal scales, and trophic levels is notoriously difficult [47, 59, 84]. Although a specialist can be defined as a species which utilises a small part of an available resource spectrum, it is rarely immediately obvious what that resource spectrum is, and whether resource spectra can be compared across different environments. For example, a spider in a rainforest experiences a very different environment to a polar bear in the Arctic. The 'banana' spider *Cupiennius* eats a wide range of different prey, from ants to centipedes [116], but this still represents a small part of the total resource spectrum available to a creature on this trophic level in a typical rainforest. A polar bear eats a monotonous diet of majoritively seals, and occasionally birds

and reindeer [76], but these constitute a very large part of the available resource spectrum to a creature on this top-predator trophic level in the Arctic. It is therefore not reasonable to compare these two species.

However, species experiencing the same, or similar, environments can be compared more usefully. Generally, specialisation of species is quantified by first isolating a group of species on the same trophic level which compete for similar resources. These resources may be different habitat, food types, or host species. Specialists can then be defined as those species which utilise a narrow range of resources, and generalists as those which utilise a wide range of resources [82]. For example, Julliard et al [82] defined specialism and generalism in a range of bird species by measuring the variety of habitat that each bird was found in. Additionally, a study by Nentwig [116] defined specialisation in non-webbuilding spiders by collecting a range of such spiders and offering a variety of prey taxa in the lab. Specialists in this case were defined as the spiders that accepted a narrow range of prey, while generalists accepted a wide range of prey.

However, it is important to recognise that although specialists and generalists are referred to as different, distinct things, species exist along a continuum of specialism to generalism [60, 89]. This continuum of specialism and generalism has been found in pollinator-plant networks [8].

In the experimental ecosystems, which we analyse in Chapter 5 of this thesis, specialists and generalists are defined by the range of environmental conditions they can thrive in. Generalists may thrive in many conditions, while specialists are restricted to only a narrow range.

1.3.3 Coexistence of specialist and generalist species

In many communities, specialist and generalist species coexist but the relative abundance of specialists and generalists (the specialist-generalist balance) may be different in different communities [34, 42, 72, 81, 84, 131]. Some studies have found coexistence between specialist and generalist species in the same community and geographical area [85, 148], while others have found that the niche breadths of species in the same geographical area tend to be similar [46, 82]. It is therefore an interesting question to ask what factors affect the specialist generalist balance in different environments, and whether the balance can be predicted by considering these factors.

An ultimate aim of this project is to be able to predict what specialist-generalist balance will be found in a particular community in a particular environment, by considering the features of the environment.

1.4 Environments, resource spectra and their effect on specialists and generalists

A number of environmental factors have been found to affect the specialist generalist balance of a community. These include the size of the habitat, its heterogeneity, the fragmentation of the available habitat and the ability of species to disperse through it. It is important to note, however, that an environment is also shaped thoroughly by the species that live there. In some cases it is easy to divide resident species into 'environment shapers' and 'competitors'. For example, when studying animals that live in trees, the tree species can be considered to be a type of environmental resource and the animals are the competitors of interest. However, in other cases the boundaries are less clear. For example, 'ecosystem engineers', such as beavers, are defined as species that change the physical landscape of the environment [80][175]. However, for the purposes of this work, we will assume that species present in an environment do not affect the physical environment, and we will only focus on studies which can clearly be viewed in this way.

The 'environment' experienced by a species is highly subjective and depends on the species in question. It depends strongly on the size of the species, its ability to disperse either itself (in the case of animals) or its progeny (in the case of plants, which cannot move). For example, a herd of elephants that may travel very large distances will experience a very different environment to that experienced by insects that live in the grass by the elephant's feet. Even if more similar animals are compared, such as lions and hyenas, they too may have widely differing territory ranges; a lion territory is of the order of 300km^2, while a hyena territory is of the order of 40km^2 [98].

In addition, resources or habitat types which are very important to one guild of species may be completely unimportant to another guild. To insect pollinators such as bees and butterflies which eat nectar, the flowers found in an environment are of key importance. However, for insects such as termites, which do not eat

nectar, the presence of flowers is irrelevant and the presence of fungi or decayed wood is important [118].

It is therefore more useful to talk about the resources experienced by a species or guild of similar species. If these resources are qualitatively similar, it may be possible to arrange them along a single axis, in which case they form a resource spectrum. This then forms a niche axis for that species, as discussed in section 1.2.1. Realistically it may not be possible to arrange all possible resources that can be exploited by a particular species along a single axis; omnivores such as wild boar may eat a wide variety of foods, including flowers, roots and large invertebrates [144] which are clearly incomparable. However, there are many cases in which resources can be arranged along a single axis, such as in the case of seed sizes for Galápagos finches [166], or in the case of the light spectrum utilised by picocyanobacteria [150]. In our modelling work, we assume all species utilise resources along a single axis.

We will now outline a number of ways in which the environment, or resource spectrum, can vary and establish the ways in which this can affect the specialist generalist balance in a community that lives in that environment.

Distribution of resources along a resource spectrum

Another potentially important factor that may affect the specialist-generalist balance is how resources are distributed along a single niche axis [e.g. food size, 102] in an environment. The way in which resources are distributed along a niche axis has not been widely studied as a whole, although some studies have focused on the effects of a particular resource spectrum on a particular system. Previous studies include the spectrum of light frequencies available for plankton growth, which is predicted to change from a smooth frequency distribution to a distribution containing spectral gaps, with increasing water depth [150], and the famous case of a seed size distribution, previously mentioned [166]. However, it is reasonable to assume that all niche axes have resources distributed along them in different ways, and that this would have some effect on the species that utilise them.

1.4.1 Homogeneity and heterogeneity

A resource spectrum may be heterogeneous, containing many different types of resource, or homogeneous containing one type of resource. The heterogeneity of a resource spectrum is known to have a large effect on the specialist-generalist balance of a community living there, with homogeneity generally favouring specialists, and heterogeneity generally favouring generalists. Heterogeneity may refer to spatial heterogeneity, or temporal heterogeneity (how often the environment changes).

Homogeneous environments favour specialists

Consider an environment with resource spectrum which is completely homogeneous, and does not change with time. For a species in such a situation, the best strategy is to specialise heavily on utilising the single resource present at the cost of utilising other resources which are not available, in order to maximise success. In addition, if the environment is competitive, any species which does not so specialise will inevitably be outcompeted by one that does.

We will use the story of bamboo and the giant panda to illustrate this mechanism and its effects. In the past, large areas of China were made up of unbroken perennial bamboo forests, reaching as far as Burma and northern Vietnam [105]. In a situation of such wide abundance of a particular reliable food source, the highly specialist giant panda evolved to eat it, and nothing else. This occurred despite the giant panda not being a 'natural' herbivore; as a member of the bear family, it possesses a single stomach more similar to carnivorous species such as dogs or cats rather than a compartmentalised foregut as found in herbivores such as cows [48]. However, with a diet so heavy in bamboo, it is thought pandas lost the gene responsible for umami taste [178], leaving the panda disinclined to eat meat. This resulted in pandas needing to consume huge quantities of bamboo to satisfy their energy requirements. In other words, the specialist giant pandas evolved to require very specific environmental conditions under conditions of homogeneity in both space and time. When these conditions changed, with the increase in farming and habitat fragmentation in more recent times, the pandas were put at a huge disadvantage, and were unable to adapt, resulting in the endangerment of pandas that we see today [105].

Homogeneous environments favouring specialists that suffer a fitness cost in other

environments has also been shown in selection experiments by Reboud and Bell [128]. In this study, unicellular algae, *Chlamydomonas*, were cultured for many generations in either dark or light conditions. The dark-adapted lineage was then placed in a light environment and vice versa, and it was found that lineages were less fit in their new environment than before they had become adapted to their old environment – showing that their niche breadth had shrunk and they had become specialised.

Temporally heterogeneous environments

Temporally heterogeneous (time-varying) environments are generally believed to favour generalists, although this depends on whether the variability is reliable (such as seasonal changes) or unpredictable (such as forest fires).

Predictable environments may allow for multiple specialist strategies in the same species (for example, trees shedding their leaves for winter, and growing them again when spring reliably comes again). Unpredictable environments generally result in generalists, although it is possible for specialists to survive in such an environment by using the strategy of 'bet hedging', in which different phenotypes of the same species are adapted to different possible environmental conditions, such that a few lucky ones will survive. This has been shown theoretically [146] and in experiments using micro-organisms [5].

Returning to the study by Reboud and Bell [128], *Chlamydomonas* that were cultured in environments that switched between dark and light every growth cycle evolved to become generalists that could survive equally well in both light and dark environments.

In addition, a study of arable fields by Fried et al [57] found that fields that used crop rotation contained more generalist weed species, while monoculture fields contained more specialist species.

Spatially heterogeneous environments

Spatially heterogeneous environments contain different types of resources or habitat that do not change in time. It is important to understand that 'spatially heterogeneous' in this context does not necessarily mean spatially clumped 'patches' of resources, and they may be mixed together. The selection pressures

in spatially heterogeneous environments are not entirely clear. Part of this uncertainty stems from experimental difficulty in creating such an environment in a well-controlled way.

A meta-analysis by Kassen [84] reviews a number of microbial experiments testing selection of specialists and generalists in spatially and temporally varying environments. This analysis concludes that although spatially heterogeneous environments are more likely to result in specialists surviving than in temporally heterogeneous environments, spatially heterogeneous environments are still more likely to favour generalists.

Fine or coarse graining of a resource spectrum

The resource spectrum, as experienced by an individual living in it, can be described as fine or coarse grained. This dichotomy is similar to that of a heterogeneous or homogeneous environment, but graining refers more specifically to heterogeneity experienced by a specific species along a single resource axis.

The grain does not necessarily correspond to the spatial variation of the landscape of the environment itself; rather, it refers to the variety of resources encountered by a particular individual throughout its lifetime [6, 55, 61, 93]. A large variety corresponds to a fine graining, and a small variety corresponds to a coarse graining. As we have learned, a species experiences a number of different resource spectrum axes, and may experience a different level of graining on each one. Fine and coarse graining on the same resource spectrum can be compared. For example, a mouse living in an arable field may experience a more coarse-grained environment then a mouse living in a wildflower meadow [153].

1.4.2 Ecotones and ecoclines

An ecotone is a sharp boundary between different and relatively homogeneous environments [63, 130, 161]. An ecocline is a more gentle ecological gradient between different types of habitat, and is relatively heterogeneous [161].

In the modelling work of this thesis, we question whether environments being ecotonal (containing large spatial discontinuities in habitat type) or ecoclinal (habitat type varies smoothly over space) has an effect on the specialist-generalist balance. This is a question not widely considered in the ecological

literature, although some empirical studies have attempted to find species that are specifically adapted to live on the ecotone. Studies of birds [14] and beetles [68] specifically find no evidence of these ecotone specialists when studying these environments.

The question of whether the presence of ecotones in a wider environment affects the specialist-generalist balance on either side of the ecotone has not been widely addressed.

Ecotones and ecoclines may also be viewed in terms of coarse and fine environmental graining. An ecotonal landscape could therefore be viewed as coarse-grained, and an ecoclinal one could be viewed as fine-grained, and predictions about specialists and generalists should therefore translate into such cases, although this has not been much explored in theoretical work.

1.4.3 Dispersal of species

Dispersal refers to the movement a species may have in space, and may occur by different mechanisms in different systems. In the case of plants, dispersal happens in the form of seeds, which may establish themselves near or far from the parent plant. In the case of animals, species may stay within a small area, or go further away to find food. In both cases, however, there are many advantages and some disadvantages when offspring disperse far from their parent. Species which disperse widely from their parent may avoid competition with kin and inbreeding. In a changing landscape, the ability to disperse may also be crucial to finding new suitable habitats, or to ease overcrowding in an area [24]. However, dispersing also carries the risk that the new habitat may be unsuitable for the species. In addition, the ability to disperse a long distance often incurs a fitness cost, resulting in lower competitive ability. This trade-off has been seen in bacterial biofilms [114], crickets [109] and plants [129].

The relationship between dispersal range and the coexistence of specialists and generalists has been investigated in previous work, but a consensus has not been reached on the impact of dispersal. Several model systems involving two habitat types have shown that low dispersal favours specialist species [26, 43, 123]. However, a recent individual-based model by Büchi et al [28] found the opposite; that specialists are favoured when dispersal rates are high. A study by Ronce and Kirkpatrick [132] finds that high and low dispersal rates favour in generalist

species, while intermediate dispersal favours specialists.

The reason for such varying conclusions may be that high and low dispersal rates have conflicting effects on specialists. Specialists may require a high dispersal range in order to find their ideal habitat; however, once they are there, philopatry[3] may be favoured, resulting in low dispersal being advantageous. The effect of dispersal range on specialists' success should therefore depend on the type of environment in question; whether it contains similar resources clustered close together, and whether it is temporally stable. The latter cases should favour low dispersal of specialists. A study of spiders by Bonte et al [21] supports the view of low-dispersal specialists, finding that 'ballooning' (the dispersal technique used by non-winged arthropods in which a 'sail' is constructed from silk, allowing them to be transported by air currents) is used far more often by generalist spiders that are found in a wide range of habitats, than by specialists that are found in few.

In Chapter 4, we will include spatial structure of environments in our theoretical model, and investigate the effects of limiting dispersal of species on the specialist-generalist balance.

1.5 Modelling ecosystems

This thesis focuses on model ecosystems, rather than empirical studies of natural environments. Modelling of ecosystems can be done in two ways: experimentally or theoretically, both of which are explored in this thesis.

1.5.1 Experimental models

Experimental models involve creating a small, self contained ecosystem, often to address a specific ecological question. Model ecosystems often use microbial life, as microbial models offer a high degree of experimental control and microbes have short generation times, making them highly suitable in particular for experiments addressing evolution. In addition, microorganisms can be genetically manipulated so that otherwise identical individuals can have well-defined genetic differences that allow specific functions to be explored [78]. For example, in the study by Dykhuizen and Davies [49] discussed in Section 1.3.1, the gene allowing

[3]'Philopatric' species have a tendency to return to or stay in the same geographical area

utilisation of lactose was deleted in the bacterium *E coli* to produce a strain that specialised on maltose. Such systems generally do not attempt to recreate the huge complexity found in nature at a small scale; rather, they allow isolated ecological hypotheses to be experimentally tested.

Earlier in this chapter we have detailed various model microbial ecosystem experiments which have explored hypotheses in ecology; for example, the study by Reboud and Bell [128] which used light and dark adapted *Chladydomonas* to explore fitness costs of adaptation to a particular environment.

As it is possible to easily create many replicates of the same system, microbial ecosystems are also useful for investigating questions about ecosystem predictability; to what extent the state of an ecosystem may be predicted, and to what extent it is dependent on stochastic intrinsic or extrinsic events. Recently, Hextra and Liebler [67] explore stationarity of microbial ecosystems by creating many replicates of a simple three-species closed microbial system. They find that fluctuations in population in the three species can be described by quantitative laws and that variations between the three species are correlated, suggesting that there may be underlying quantitative laws that apply across different types of ecosystem that may be used to predict events, or at least to predict the relative frequency of events.

Macroscopic ecosystems may also be experimentally modelled. For example, as discussed previously in Section 1.1.4, a number of lakes in Canada were manipulated a 37 year experiment to determine causes of eutrophication in lakes by adding fertiliser to them in varying ratios of nitrogen and phosphorous [139], resulting in the conclusion that eutrophication of lakes cannot be controlled by reducing nitrogen input.

The final chapter in this thesis details two microbial model systems which we use to firstly investigate the effects of varying sulphate and cellulose concentrations on a community sampled from a freshwater pond, and secondly to investigate alternative ecosystem states that may arise in communities sampled from the same pond.

1.5.2 Theoretical models

The main focus of this thesis is on theoretical models for ecosystems. Theoretical models can be classified into analytic models and computational models, the latter of which may be equation or agent based. Historically, analytic models were used to understand ecosystems, but with the recent increase in computing power, more complex computational simulations can now be achieved. In the remainder of this section, we will outline some different methods in modelling ecosystems, focusing particularly on those which model competition between species.

1.5.3 Lotka-Volterra competition models

Lotka-Volterra models were introduce to study two species interacting with a predator-prey relationship. These models are well known for producing the famous staggered oscillations that are characteristic of many predator-prey populations cycles seen in the natural world [102, 165].

However, the same type of equations can be used to model competition between two organisms that are competing for the same resource. In this context the Lotka-Volterra equations take the following form,

$$\begin{aligned} \frac{dN_1}{dt} &= r_1 N_1 \left(1 - \frac{N_1}{K_1} - b_{12}\frac{N_2}{K_1}\right) \\ \frac{dN_2}{dt} &= r_2 N_2 \left(1 - \frac{N_2}{K_2} - b_{21}\frac{N_1}{K_2}\right) \end{aligned} \tag{1.6}$$

where N_1 and N_2 are the abundances of species 1 and 2, b_{21} and b_{12} are coefficients that describe how much the two species compete with each other, r_1 and r_2 are the birth rates of species 1 and species 2, and K_1 and K_2 are the carrying capacities of species 1 and species 2 [113][165].

Equation 1.6 predicts that if b_{12} is much larger than b_{21}, then the presence of species 2 has a much larger negative effect on species 1 than vice versa, and so species 1 will be outcompeted, and eliminated. However, if b_{12} and b_{21} are small, the two species may coexist. It is possible to add more species to the model, by simply adding extra equations and competition terms. Lotka-Volterra models can be seen as a 'top down' approach to modelling ecology; only the

final outcome of competition between two species is seen as relevant, while the underlying mechanisms for the competition are ignored.

There have been some attempts to apply Lotka-Volterra competition models to experimental data. Wilbur [170] tested the assumptions of the model in experiments with different species of mole salamander. Different species of salamander were placed in competition with one another, in groups of two and three, with the result that some species coexisted while others drove their companions to extinction. However, Lotka-Volterra models proved to be inadequate for interpreting natural communities, as they did not predict experimental results.

Analytical solutions for multi-species Lotka-Volterra models have been found [53] but in such studies the focus is primary on the mathematics, while relevance to real ecosystems is generally not considered. In a review of the subject, Tilman [155] notes that a multi-species community is extremely complex, and competitive interactions between species may become completely unpredictable. In attempting to build a multi-species model of competition that has any relation to a natural ecosystem, many coefficients between species b_{ij} must be experimentally or empirically found. Bender [19] introduces a possible way of doing this, called a 'PRESS' experiment. In such an experiment, a community of species is allowed to come to an equilibrium, and then the population of each species is individually perturbed, usually to zero, to determine the immediate effect on the remaining species. However, doing such an experiment with a real ecosystem is completely impractical, and the removal of a single species in such an experiment might completely change the results. Thus, although the top-down approach that Lotka-Volterra models offer seems as if it may simplify the problem of competition, in a many-species situation it becomes hopelessly complex.

Although using a Lotka-Volterra model to model a community of species explicitly is not practical, they may be used to model a hypothetical community of species, with the aim of deducing some information about community interactions in general. For example, Coyte et al [39] use a Lotka-Volterra model of a hypothetical microbial network with both competitive and cooperative links to show that co-operative links in a community create instability in the network, while competitive links stablise. Yoshino et al [177] use a Lotka-Volterra model to investigate interactions in food webs, finding that the presence of higher order interactions between species lead to greater diversity when the system is at stationarity.

1.5.4 Individual based models

In contrast to Lotka-Volterra models, which take a top down approach, individual based models take a bottom-up, mechanistic approach to representing an ecosystem. Here, individuals of a community are modelled as independent agents, and the overall behaviour of the community emerges as a result of many independent interactions between the agents.

With the ongoing increase in computing power, use of individual-based models has risen considerably in recent years and they are now used in almost all branches of ecology, especially in cases where a spatially explicit model is needed [44].

Individual based models can generally be divided into two categories: highly specific 'tactical models' which attempt to make quantitative predictions about a real system, and more abstract 'strategic models' which attempt to make qualitative predictions of general ecological systems [51].

By using tactical models, numerous successful predictions of animal behaviour have been made [147]. For example, inSALMO, an individual based model designed to predict habitat effects on salmon [127] successfully predicted the movements of salmon at different life stages in a Californian spawning stream.

In contrast, strategic models do not attempt to make such specific predictions, but rather aim to uncover more general effects, such as the effect of competition and mutualism between hypothetical organisms under the effects of climate change [25]. The computational model we present in Chapters 2, 3 and 4 of this thesis falls into this category.

1.5.5 Neutral models

Lotka-Volterra and individual based models can both be thought of as 'niche' models, which assume that processes in ecosystems must be highly influenced by differences between species. Where niche models take careful account of what each species is doing, and how they interact with one another, neutral models assume that all species behave exactly the same and community assembly is a result simply of stochastic processes.

Neutral models have been used in population genetics for many years; examples include the Moran model and Wright-Fisher model [86]. They are attractively

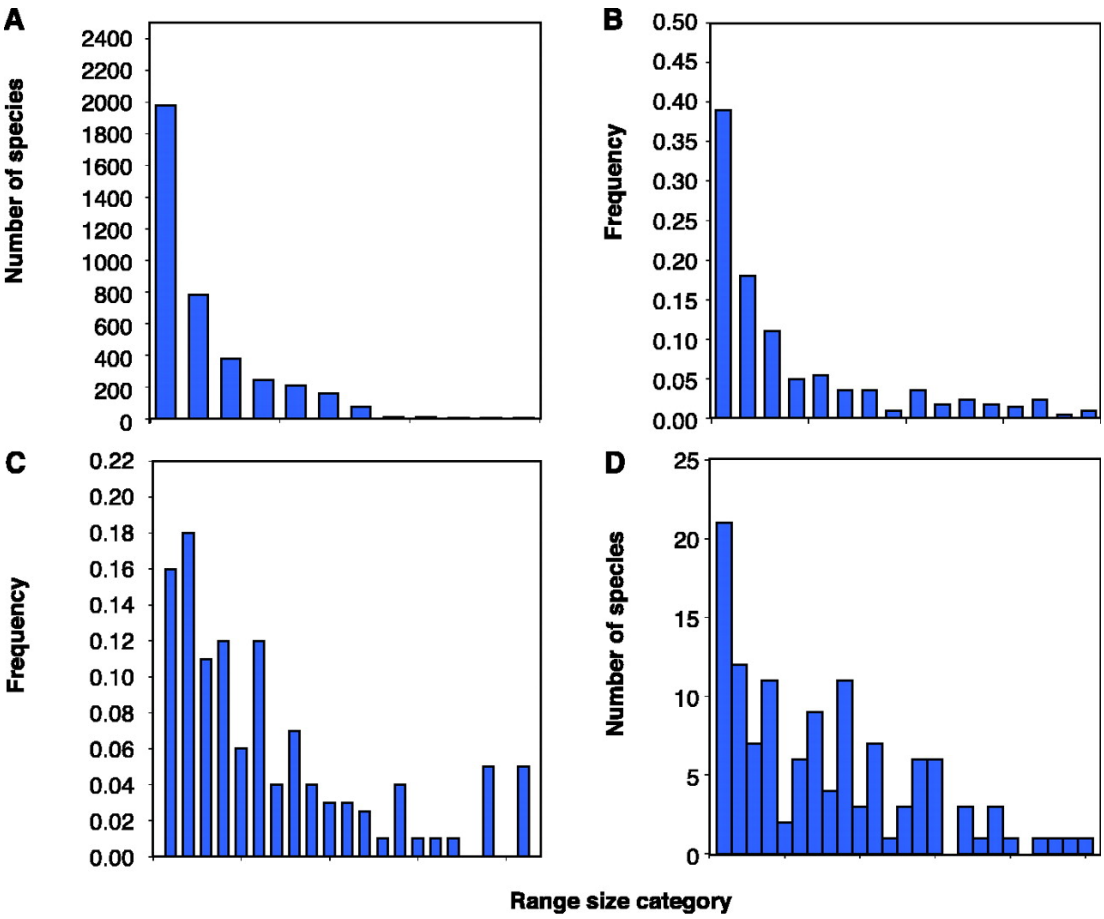


Figure 1.9 *A set of abundance distributions from (a) New World birds, (b) passerine birds in Australia, (c) North American birds, (d) a neutral model. (Adapted from [18])*

simple, and analytical results can be obtained with only a few parameters. However, their use in ecology to predict community patterns is more recent and more controversial, possibly due to the fact that different species in a real ecosystem are plainly not identical [18]. Nevertheless, highly influential work originated by Hubbell [71] aims to take neutral theory and establish it as a central tool in ecology theory.

The basic premise of a neutral model is simple. We assume a number of trophically similar species in the same area that are competing for the same resources. Usually, neutral models are assumed to have a finite population size, although this is not crucial.

The probability of any individual of any species reproducing or dying is assumed to be constant, meaning that a species population may increase or decrease with equal probability. Every individual in the system has an equal probability of

either reproducing itself or dying. The population of the species as a whole is simply the sum of all of its individuals; the abundance of different species is therefore entirely stochastic. Some species are fortunate, reproduce many times, and become highly abundant. Others simply die out, and are replaced by immigrants [71, 96].

For many years, similar patterns of abundance have been seen in very different ecological communities and ecologists aimed to explain such patterns in terms of niche theory, suggesting perhaps that each ecological community, no matter how big or small, had particular roles that were fulfilled by a different species in each one [18]. However, it was found that neutral models produced abundance patterns that were indistinguishable from the empirically observed patterns found in natural communities. Figure 1.9 shows three abundance patterns for real world bird communities, compared to similar data generated by a neutral model.

The correct interpretation of the success of neutral models in reproducing abundance patterns so effectively is debated. Some argue that the success of the neutral model simply shows that abundance patterns are not a useful thing to measure when assessing a community and that the neutral model should simply be taken as a 'null' model. To accept the existence of neutral dynamics per se would be in complete contradiction of the principle of competitive exclusion, which has been held as central in ecology for many years [90].

Others argue that neutral dynamics truly do have a huge impact on community assembly [18]. A third interpretation, seeking to unite the two paradigms of niche and neutral mechanisms, states that different communities exist along a spectrum of niche and neutral. In other words, some highly specialised communities interact according to niche dynamics, while other more generalised communities are governed by neutral dynamics [70, 90]

Niche-neutral continuum models

Neutral theory takes the extreme approach of assuming that all species are identical and the only processes taking place are stochastic. Traditional niche theory, by contrast, assumes that there are no stochastic processes happening at all. Recently, models have been developed which attempt to unite the two approaches. These models attempt to explain how species with explicit niche differences can nonetheless result in a community that behaves according to

neutral dynamics [15, 64, 83, 117, 142, 156], and also investigate how different factors affect where a community sits somewhere along a spectrum between niche and neutral behaviour.

Of these models, many (including the models used in this thesis) take the approach of using an agent-based simulation in which many species compete with each other for sites on a grid. Each grid square has a particular 'resource value', and each species has a unique response to the local resources which is typically modelled on a Gaussian function which depends on each species 'niche ideal', i.e. ideal resource value, and 'niche breadth' (i.e. acceptable range of resource values). In these models the competition has an element of randomness, as in the neutral model, but probabilities of success for a given species are weighted according to how adapted a competing species may be to an available site, bringing in an element of niche theory. Using models of this sort it has been shown [15, 64] that species that have explicit niche differences can behave according to the predictions of neutral dynamics.

In addition, different possible inputs and their effect on the resulting niche-neutrality of the resulting community can be explored. For example, Gravel et al [64] found that a high immigration rate resulted in a community exhibiting more neutral behaviour, while a low immigration rate resulted in a community where species were regularly spaced along the niche axis, which is usually a feature of a community subject to niche dynamics. More recently, Bar-Massada et al [15] found that increasing the heterogeneity of the environment resulted in the community behaving less neutrally.

The computational model we develop in this thesis is very similar to these niche-neutral continuum models, but by varying the niche breadth we have used it to instead explore competition between specialists and generalists. In agreement with earlier studies, our results suggest that different species may be behaving according to more niche or neutral dynamics within a single community.

1.5.6 Modelling specialists and generalists

The focus of this thesis is on competition between specialist and generalist species. Here, we outline some previous models attempting to explore the same thing.

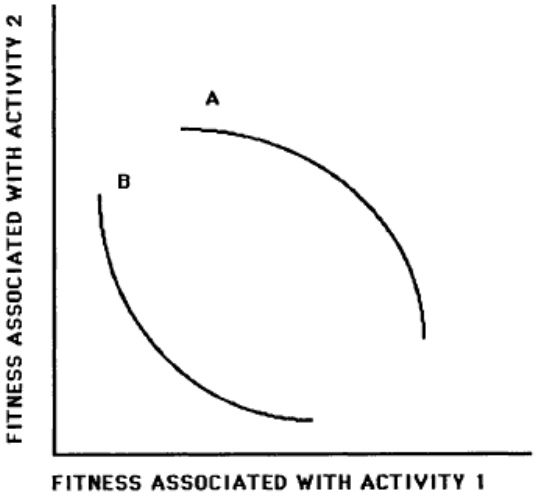


Figure 1.10 *Concave (B) and convex (A) fitness sets, favouring specialists and generalists respectively (Reproduced from [171])*

Early analytic models

One of the first attempts to model the effect of the environment on the specialism strategies of resulting species was made by MacArthur and Pianka [97] in 1966. They used a simple analytic model, based on the principles that a predator must use as little energy as possible to catch its prey in a given environment, to show that large homogeneous environmental patches should result in independent specialist populations, while smaller patches, resulting in a more heterogeneous environment, should result in a generalist population, assuming a jack-of-all-trades trade-off.

Modelling attempts over the next thirty years tended to focus on a small number of species, usually two specialists and one generalist, using equations that could be solved analytically. Wilson and Yoshimura (1994) [171] review these models. Taken together, they found that coexistence between specialists and generalists depended on a number of factors.

The first of these factors is the shape of the fitness set, an idea developed by Levins [92], which represents the severity of the trade-off between specialising on a single resource, or generalising on both. Figure 1.10 shows two possible fitness sets. If the fitness set is concave (B), specialisation is predicted to evolve, as there is a large cost associated with being fit at both activities. If it is convex (A), generalism is predicted to evolve. A number of models conclude that the fitness set must be concave for specialisation to occur [171].

The second factor is activity selection, which refers to whether species can choose to perform only a few activities or are forced to perform all activities, even if they are performed badly. Specialisation is predicted to be favoured when species can choose their activities, as they can then only perform activities they are good at, and avoid those they are not [171].

Thirdly, these models predict that negative density dependence can affect specialist-generalist balance. This refers to the fitness cost that may occur when a species attempts to perform an activity already being performed by many other species. In particular, specialists are more likely to succeed when attempting to fill an empty niche [171]. Considering the case of the tree dwelling lizards discussed before [140] in Section 1.2.1, it seems likely that this is an active factor in specialisation.

Finally, as we have discussed before, temporally varying environments are predicted to affect specialisation, while high temporal variability is predicted to result in generalists [171].

Recent individual based models

More recently, the balance between specialists and generalists has been investigated using more complex, multi-species models.

Nagelkerke and Menken [115] used a multi-species, multi-habitat model to show that the size and number of available different habitats can drastically affect the specialist generalist balance. In this study, a large number of habitats of small size favoured a single generalist winner, but if the habitats became larger, a single specialist was able to displace the generalist in each one. These results are not seen with traditional two specialist, one generalist models, as the generalist needs a wide range of habitat over which to maintain its advantage.

Similarly, Weiner and Xiao [167] used a multi-species, multi-habitat model with varying fecundity of all species to show that when fecundity is low, the generalist is predicted to win, while an intermediate fecundity leads to specialist and generalist coexistence, and high fecundity leads to only specialist surviving on their ideal habitat.

Very recently, individual based models very similar in form to the niche-neutral continuum models discussed before have been used to investigate ecosystems with

many species with a spectrum of specialism and generalism, interacting in an environment with continuous habitat type. In particular, Büchi et al [28] used such a model, in which the niche breadth of species varies continuously and dispersal is limited, to find that when dispersal is highly limited, generalists are predicted to evolve, and when species may disperse a long distance, specialists are predicted to evolve that may disperse to find their optimum habitat. In Chapters 2, 3 and 4 of this thesis, we use a closely related model to investigate the distribution of the resource spectrum on the specialist generalist balance.

In Chapter 2, we introduce our agent-based model of specialist and generalist competition and show that it reproduces a number of phenomenon seen both empirically and in previous models and explore which factors lead to specialist and generalist coexistence.

In Chapter 3, we focus on the effects of the size of discontinuities and grains in the resource spectrum and introduce a theory which we term 'resource spectrum engineering', in which extreme specialists occupying favourable sites create effective gaps in the resource spectrum for less specialised species, giving more mild specialists an advantage over generalists.

In Chapter 4, we introduce space to our model by limiting the dispersal of species. We explore the relationship between dispersal range and specialisation, and establish how different arrangements of resources in space affect the specialist-generalist balance.

In Chapter 5, we turn to analysing microbial sequence data from two experiments of microbial communities. In the first experiment, samples of water and sediment are taken from a freshwater pond, and placed into microcosms with different concentrations of added cellulose and sulphide. From this, we can speculate which species are generalist (thrive in a wide range of environmental conditions) and which are specialist (thrive in only a few). In the second experiment, many replicate microcosms are used to explore alternative ecosystem states. We use network analysis to find possible interactions between species, and note that generalist species may interact more than specialist species.

Chapter 2

An agent-based model of a competitive environment

This chapter will provide details of a stochastic, agent-based simulation model that we have used to explore the effects of competition on community assembly, and builds on previous work by Gravel et al (2006) [64] and Büchi et al (2014) [28]. To explain very briefly, this model puts different species in competition with one another for resources. In the full model, the efficacy of a species at winning resources is dependent on three life history traits: its desired type of resource (μ), its niche breadth (degree of specialism or generalism, σ) and its ability to disperse its offspring to more distantly located resources (δ). If σ is set to a maximum, the resultant species is blind to the type of resource it consumes. If δ is set to a maximum, the species can disperse anywhere in the environment with equal ease. If the δ values of all species are set to a maximum, the model becomes well-mixed.

The niche breadth σ can either be held constant, allowing us to explore the behaviour of a community of generalists or specialists, or allowed to vary, allowing exploration of competition between generalist and specialist strategies. Similarly, by holding the dispersal parameter δ constant, we can explore the effects of an environment with limited dispersal on community composition. By allowing δ to vary, we can explore the competition between long-range dispersal strategies and short-range ones.

Key to the outcome of these simulations is the environment within which the community competes. The nature of the environment depends on two things; the

underlying resource spectrum, i.e. which resources exist in the simulation, and how these resources are arranged in space. If the simulation is well-mixed, then the latter does not apply.

2.1 Full model description

In this model, species compete for resources on sites in an environment. The environment is represented by an X by X 2D lattice, in which $X \times X = N$, where N is the total number of sites. In all of the simulations described in this thesis, $X = 100$ and $N = 10000$.

Each site on the lattice is assumed to contain one environmental resource, described by a parameter E that can be utilised for one time step by one individual of a species that is located at that site, after which all resources are replenished. We assume that each site is always occupied by one individual, resulting in a constant N individuals in the environment at all times.

Each E value takes the form of a number between 0 and 1. This can be thought of as describing a range of environmental conditions, for example, seed sizes, or salinity concentrations. The boundaries of E are periodic; for example, a site with $E = 0.95$ and a site with $E = 0.05$ are considered similar rather than dissimilar. These periodic boundaries are clearly not realistic, but making the E axis periodic avoids behaviour resultant from edge effects and does not ultimately change our results.

This chapter and the following chapter focus on models in which dispersal is not limited ($\delta \to \infty$) and so the model is well-mixed. In a well-mixed model, the arrangement of sites in space is not important. For this reason, we will leave the description of the arrangement of sites in space to Chapter 4.

In our model, a 'species' is a group of individuals that have the same life-history traits. All individuals of a species are identical to one another and reproduce asexually; all individuals therefore behave as independent agents.

Species have three life-history traits that determine their individuals' success in the model; μ, σ and δ. δ determines the dispersal distance of the species; in this chapter and the following chapter it is set to ∞ for all species but it will return to our discussion in Chapter 4.

The parameter μ describes the preferred E value of a species. When an individual is competing to win a site, its chances will be higher if its μ value is close to the E value of the site it is competing for. Like E, the μ value for every species falls between 0 and 1. When a new species is created, its μ value is randomly picked from a uniform distribution between 0 and 1.

The final life history trait, σ, is the niche breadth of a species; this describes how good it is at competing for sites with E values close to and far from its μ value. A high σ means the species is a generalist (can use a wide range of resources) and a low σ means it is a specialist (can only use a narrow range of resources). When a species is created, its σ value is created by randomly picking from a uniform distribution between -3 and 1, and then raising 10 to the power of that number. This results in a uniform distribution in log space between 0.001 and 10. A species with σ=10 can compete equally well for all sites in the environment (i.e., its μ value becomes irrelevant, because its E value is much larger than the range of E values present in the model).

In each time step, we assume that all individuals in the environment die. The newly vacant sites are then occupied either by a copy of an individual from the previous generation, or by an immigrant. An immigrant appears with probability I/N, such that I is the number of immigrants per time step. In all of our simulations, I=20. An immigrant is assumed to be a new species, created using the same rules as for the species created at the beginning of the simulation. This immigration rate allows the environment to be exposed to a wide range of strategies, which is necessary to allow for the establishment of highly specialised species.

This process of repopulating the lattice at every step can be thought of as all individuals releasing $N - I$ juveniles which disperse to every site (apart from the ones selected to receive immigrants) such that every site initially contains $N - I$ juveniles. These juveniles then compete with one another to become established on that site and grow to adulthood. A juvenile i is selected with a probability proportional to its fitness at that site, which is computed by its 'juvenile survival function', which is dependent on its life-history traits.

Computational algorithm

Computationally, this process is carried out using the following steps in the well-mixed case:

1. When a species is created, the fecundity function for that particular species is evaluated for all sites in the lattice and stored for future use.
2. At the beginning of each time step, a copy is made of the state of individuals present in the lattice. This copy is then held static until the beginning of the next time step.
3. The number of individuals of each species is added up and stored.
4. Every site on the lattice is then visited. The following steps happen at each site:
   (a) The individual occupying the site from the previous generation is removed.
   (b) Each individual 'competes' to be the one whose offspring establishes at the newly vacant site. A weighted probability of establishment is calculated for each species. This is given by the following equation, where $F(\mu, \sigma; E)$ refers to the juvenile survival function:

$$\text{Weighted prob} = \text{No. of individuals} \times F(\mu, \sigma; E) \tag{2.1}$$

   (c) All weighted probabilities from all species are added up and normalised such that they sum to 1.
   (d) The following algorithm is then applied:

```
PROBS = array containing the weighted probability of each species
INDICES = array containing the identifying index of each species
A = random number between 0 and 1 picked from a uniform distribution
i = 0
COUNT = PROBS[0]
loop
    if COUNT ≥ A then
        return INDICES[i]
    else
```

```
            COUNT = COUNT + PROBS[i]
            i+=1
        end if
    end loop
```

(e) The identifying index of the winning species has been selected, and an individual of that species is placed in the site.

By using the same copy of the previous lattice state to populate the new generation, rather than the current lattice state, it is irrelevant what order sites are visited on the lattice.

To speed computation time, the evaluation of the juvenile survival function at each lattice site for each species is computed only once as the species is created and then stored for use throughout the simulation. Note that this is only possible because the environmental lattice is static and the simulation is well-mixed.

The computational complexity of this algorithm can be estimated by finding the number of calculations required to perform a single time step. One time step involves two stages - firstly, I new species must be created, and N calculations performed in order to evaluate the fecundity function at each site for each new species. Secondly, a calculation must be made for each site to establish which species succeeds in replicating itself on the site. This calculation, following the algorithm above, will on average require $\frac{N}{2}$ calculations. On average, the number of calculations needed to select a winning species is $\frac{s}{2}$, where s is the number of species, and for some parameter sets this may be equal to N. The computational complexity therefore scales with Ns.

The juvenile survival function

We assume that a trade-off exists between large and small σ values. We assume that a highly specialised species with a low σ value is very good at competing for sites with E values very close to its μ value, but essentially unable to compete for any other sites in the rest of the environment. A species with a high σ value has a chance of winning any site in the environment, but will not compete very well for any of them. This can be thought of as the 'jack of all trades is master of none' assumption. For example, the most specialist species possible with $\sigma = 10^{-3}$ will be $400\times$ as likely as the most generalist species possible with $\sigma = 10^{1}$ to win a

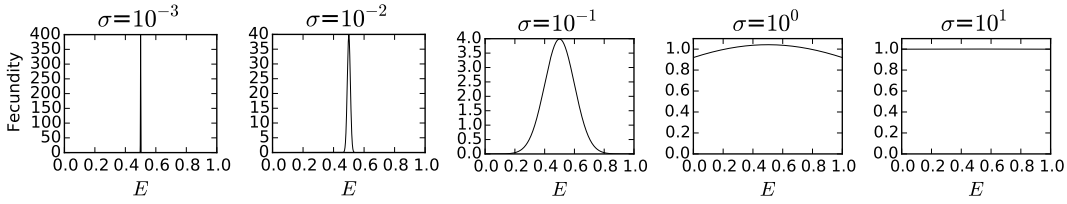


Figure 2.1 *Juvenile survival function over environmental resource E for a range of niche breadth σ, for an individual with preferred environmental value μ=0.5.*

If σ is low, a species competes very well in a very narrow environmental range, but effectively cannot compete at all in most other sites; it is therefore a specialist. If σ is high, a species competes equally well everywhere, but particularly well nowhere; it is therefore a generalist.

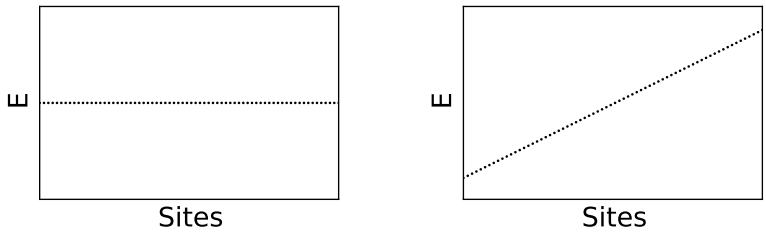
site with its ideal E value, but the generalist has a higher chance of winning the vast majority of other sites in the environment.

In the model, the competitiveness of an individual across the environment is given by computing the 'juvenile survival function' of its species. The juvenile survival function of a species i is essentially a Gaussian function with width σ_i, centred at μ_i (Equation 2.2).

$$F(\mu,\sigma;E) = \frac{\exp\left(-\frac{\Delta(\mu,E)^2}{2\sigma^2}\right)}{\int_{-\frac{1}{2}}^{\frac{1}{2}} \mathrm{d}x \exp\left(-\frac{x^2}{2\sigma^2}\right)} \qquad \Delta(\mu,E) = \begin{cases} |\mu - E|, & \text{if } |\mu - E| < 0.5 \\ 1 - |\mu - E| & \text{otherwise} \end{cases} \tag{2.2}$$

Using a Gaussian response curve carries with it many assumptions about the nature of the response and these assumptions must be considered. A Gaussian is completely symmetric, while other possible curves, such as a Beta response curve, have a degree of skew. We must therefore consider whether a symmetric or asymmetric curve is more likely to be seen in nature. Oksanen and Minchin [119] examine response curves for vascular plants along an altitude gradient and find that symmetric, bell-shaped response curves are common, concluding that 42 % of the plants investigated could have been modeled using a Gaussian response curve. We therefore conclude that using a Gaussian response curve is reasonable.

This Gaussian is normalised to incorporate the trade-off described above; this ensures a fair competition between generalists and specialists. Thus, for all species, the area underneath their juvenile survival function is set to be equal.



(a) *Homogeneous*

(b) *Uniformly heterogeneous*

Figure 2.2 *A small sample of homogeneous and uniformly heterogeneous resource spectra.*

In the homogeneous case, all environmental sites have the same E value; in the uniformly heterogeneous case, all environmental sites have a different value and are equally spaced.

The area under a Gaussian normalised in the standard way is 1, and for specialist species this holds, as their juvenile survival function does not extend far out of the bounds of E (0 to 1). However, for generalist species, their juvenile survival function is much wider. For a perfect generalist ($\sigma = 10^1$), the Gaussian is so wide such that the crest of it, falling between 0 and 1, is effectively flat. To normalise the Gaussian in the standard way would therefore put the generalists at a severe disadvantage.

We therefore normalise each species' juvenile survival function by dividing it by the area found under a Gaussian centred at 0, integrated between $\frac{1}{2}$ and $-\frac{1}{2}$. This results in the same normalisation factor for all species of the same σ value, regardless of μ. The full juvenile survival function of a species in a well-mixed system is thus given by Equation 2.2. Note that in a system in which all species are of the same σ, the normalisation is not needed.

Figure 2.1 shows the normalised juvenile survival function for a different species with a range of σ, all with μ=0.5.

2.2 Homogeneous and heterogeneous environments

A central aim of our simulations is to establish the effect of the environment that species compete on in our model on the resulting specialist-generalist balance.

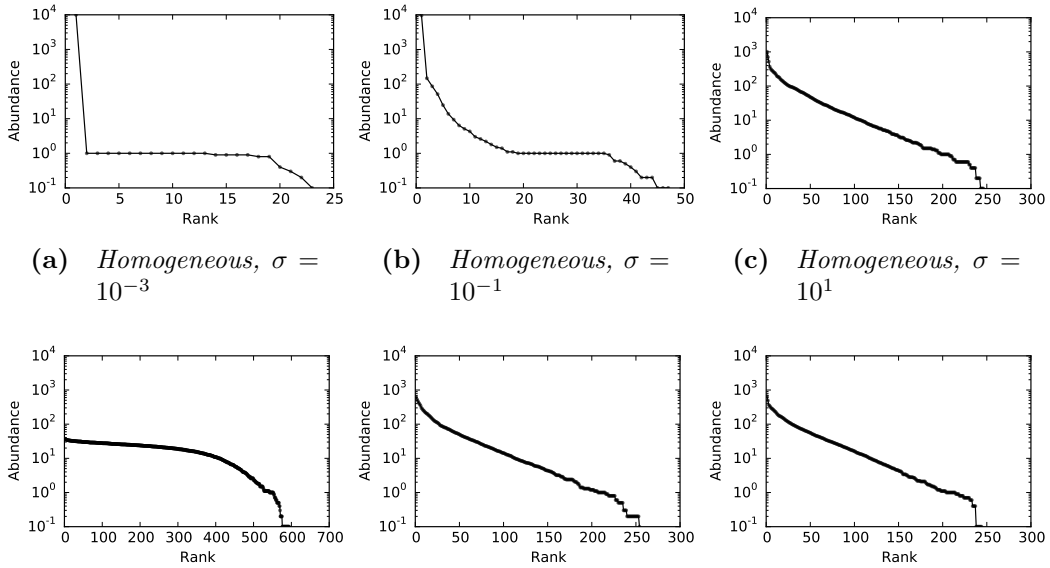
We will introduce these simulations with two simple types of environment, one completely homogeneous and one that is heterogeneous in a uniform way. Initially, all of our simulations will be well-mixed, so only the resource spectrum of available environmental sites throughout the habitat is important. Figure 2.2 shows an illustration of two simple types of resource spectrum. Figure 2.2a shows a completely homogeneous environment in which every site contains a resource with equal environmental parameter $E = 0.5$. This would correspond, for example, to an environment containing only seeds of one size. Figure 2.2b shows a 'uniformly heterogeneous' resource spectrum in which every site has a unique E value, with these values being distributed uniformly between 0 and 1, with an interval of 1/N between each one. As this model is well-mixed, the arrangement of resources in space is not important.

2.3 Results: A well-mixed environment with constant σ

To explore the behaviour of our model and to establish the ways in which we will analyse more complex models later in the thesis, we first examine simple cases, in which σ is held constant; i.e. all species have equal niche breadth σ.

We then populate the environments shown in Figure 2.2 with species of the same niche breadth σ and run the simulation for 5000 time steps. After this time, the system is in a steady state. We identify the steady state in this case by a rank abundance distribution which is unchanging for over 2000 time steps. We can then examine the resultant species present as a function of each σ value chosen.

To analyse these results, we can plot rank abundance curves (Figure 2.3). A rank abundance curve is a way of visualising the community makeup. Species are ranked by their abundance and then that abundance is plotted against rank, such that the most abundant species is of rank 1, and so on. The top line in Figure 2.3 shows rank abundance plots for the homogeneous environment, with increasing values of σ. When σ is very small, the species that is fittest (has a μ value closest to 0.5) dominates, as we see in Figure 2.3a; this species has an abundance very close to N. All of the less abundant species have abundance 1; these are newly arrived immigrants that will almost certainly not succeed in replicating. As σ increases in our simulations, this dominance by a single fittest species softens. In



(a) *Homogeneous,* $\sigma = 10^{-3}$

(b) *Homogeneous,* $\sigma = 10^{-1}$

(c) *Homogeneous,* $\sigma = 10^{1}$

(d) *Heterogeneous,* $\sigma = 10^{-3}$

(e) *Heterogeneous,* $\sigma = 10^{-1}$

(f) *Heterogeneous,* $\sigma = 10^{1}$

Figure 2.3 *Rank abundance plots for species present, for three* σ *(niche breadth) values in two types of environment.*

In the homogeneous environment, a single species that is specialised to the environment that is available can dominate, as seen in (a). In the heterogeneous case, many specialists can co-exist, each specialising on a different section of the available resources, as seen in (d). Generalists have the same abundance patterns in both types of environment (c,f).

Figure 2.3b in which $\sigma = 10^{-1}$, we see that the species with rank 1 still has an abundance close to N, but the second most abundant species has an abundance of 100, and the tail of the plot is longer. In Figure 2.3c, where the environment is now populated only by generalists of σ=10, we see a much more egalitarian rank abundance curve, in which the first and second most abundant species have abundances close to one another.

The bottom row in Figure 2.3 shows rank abundance curves for the heterogeneous environment. For the highly specialist community of Figure 2.3d ($\sigma = 10^{-3}$), we see that the most abundant 400 species have an abundance roughly equal to each other, falling between 40 and 10. This is an example of niche partitioning, a mechanism by which many specialist species may coexist beside one another, each utilising resources optimal to their specialism and not infringing on each other. In Figures 2.3e and 2.3f, we see rank abundance curves very similar to that of Figure 2.3c, although the community makeup changes slightly as we will see in Figure 2.4.

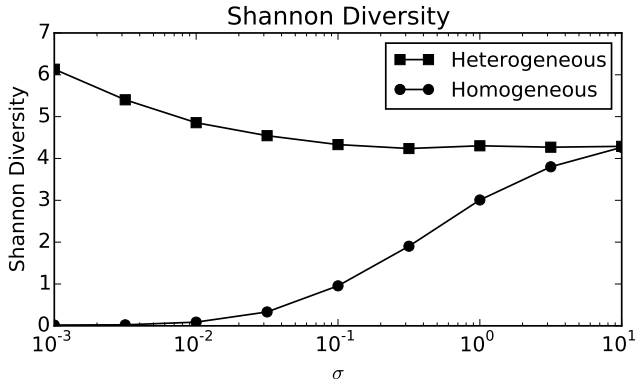


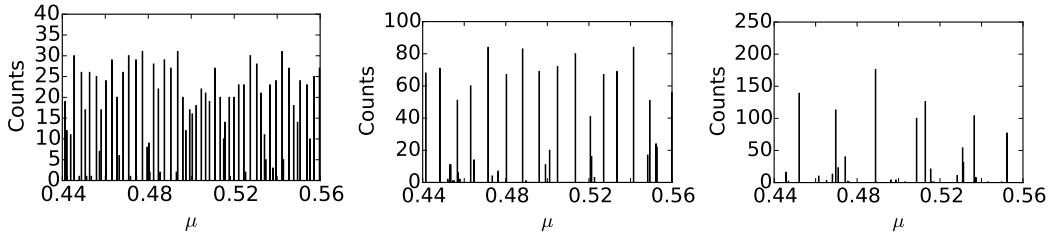
Figure 2.4 *Shannon diversity for simulations with fixed niche breadth σ, over a range of σ in two simple environments shown in Figure 2.2. The difference in effect between the two types of environment lessens as species are made more generalist.*

To quantify the diversity seen in these rank abundance curves into a single number, we can calculate the Shannon diversity of the community and plot it for a range of σ values (Figure 2.4). We can clearly see that when σ is low (meaning a specialist community), the heterogeneity of the environment has a huge effect on the overall diversity, but when σ is large (generalist community), the environment becomes irrelevant. This is because specific resources are critical to the survival of specialists, but generalists are unaffected by the type of resource they consume.

2.3.1 Regularly spaced niche partitioning

In our uniformly distributed heterogeneous environment (Figure 2.2b), how do specialists arrange themselves into niches? Our simulations reveal an interesting phenomenon. We find that putting a set of species with the same σ value together in an environment causes a regularly spaced *niche partitioning* to spontaneously emerge. Examining the μ distribution at the end of a simulation reveals this partitioning (Figure 2.5)

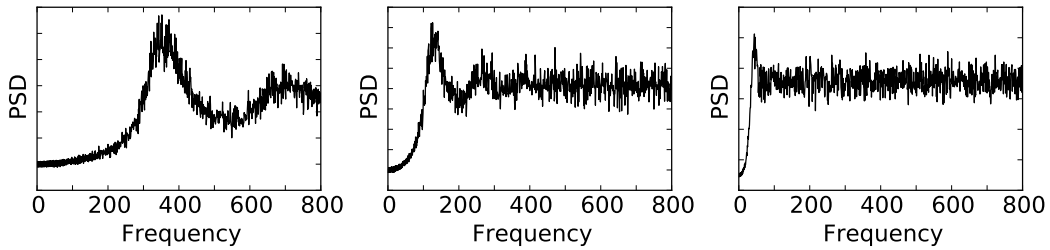
To examine how far apart in E space species partition themselves, we can take a Fourier transform of the distribution of μ (environmental preference) values, taken at the very end of the simulation (Figure 2.6). For the most specialist species ($\sigma = 10^{-3}$) we see that the peak frequency occurs at 350 species/unit. This gives a gap of roughly 0.00285 (1/350) between the μ values of each established species, although clearly as the frequency curve is somewhat spread out there are some



(a) $\sigma = 10^{-3}$ **(b)** $\sigma = 10^{-2.5}$ **(c)** $\sigma = 10^{-2}$

Figure 2.5 *Niche partitioning at low niche breadth σ.*

Surviving species spontaneously partition to minimise competition, which can be seen by the regular spacing of μ values.



(a) $\sigma = 10^{-3}$ **(b)** $\sigma = 10^{-2.5}$ **(c)** $\sigma = 10^{-2}$

Figure 2.6 *Fourier transform of niche partitioning shown in Figure 2.5 at low niche breadth σ.*

Power spectral density (PSD) against frequency. Units of PSD are arbitrary and so have not been included. We can see from this that species have partitioned themselves at regular μ intervals to minimise competition.

larger and some smaller gaps. This gives a niche spacing of around 2.85 $\times\sigma$, although there is clearly some variation as the peak in frequency is spread out. For $\sigma = 10^{-2.5}$, we see a peak at around 125 species/unit, which corresponds to a spacing of 0.008, which is around 2.53 of the σ value.

It is interesting that this niche partitioning emerges despite no species initially having an advantage. In these simulations, every species has the same σ value, and in a uniformly distributed environment, each species ought to have the same chance of survival. Despite this, some species are selected while others whose μ values fall between these peaks, do not prosper. This niche partitioning, and ultimately the species that will be triumphant, emerges initially presumably through stochastic events.

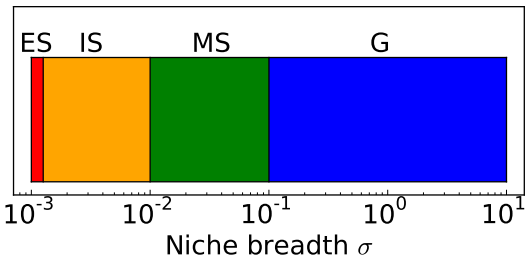


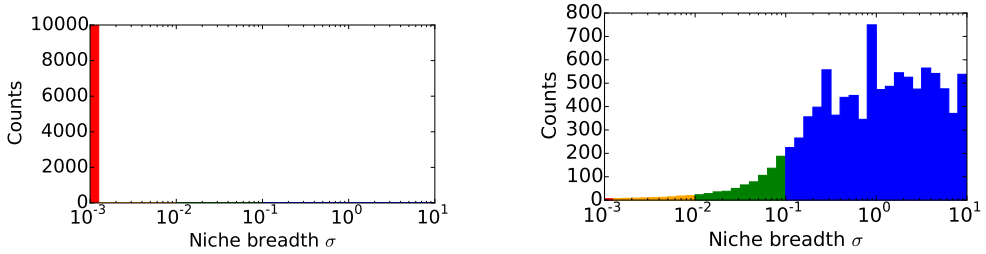
Figure 2.7 *Categories of σ*

Species will be categorised into four different groups according to their σ value. The boundaries between different σ groups are arbitrary.

2.4 Using variable σ to explore specialist generalist competition

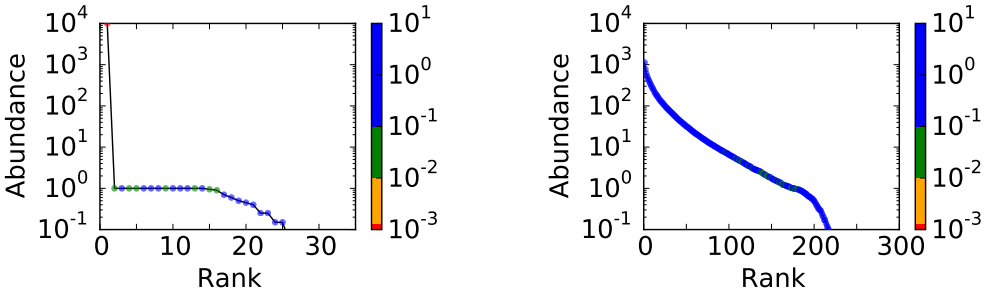
We now turn to the interesting case in which σ is allowed to vary within a simulation. This allows us to simulate the effect of generalist species competing for space with specialist species. We use the same two simple environments as before (Figure 2.2 and run simulations for 30,000 time steps, taking around twenty replicates for each parameter set. 30,000 time steps was judged to be long enough to allow the system to reach steady state in most cases, or to a state of slow relaxation in other cases in which it may be assumed that the qualitative appearance of the system will not change with additional time. This took around five hours to run, when code was written in C++.

Although our simulations allow species to take any value of σ between the range of 10^{-3} to 10, for easier analysis of our results we categorise species into four types according to their σ values: Extreme specialists (ES), Intermediate specialists (IS), Mild specialists (MS) and Generalists (G). These are colour-coded as in Figure 2.7; this colour coding will be used in our plots throughout the thesis. Note that the categories do not all have the same size of σ-range; in particular extreme specialists account for only a small percentage of the total strategies available. The system was judged to have reached steady state when the proportions of these σ categories were unchanging for over 5,000 time steps.



(a) *Homogeneous environment, σ distribution*

(b) *Heterogeneous environment, σ distribution*



(c) *Homogeneous environment, rank abundance*

(d) *Heterogeneous environment, rank abundance*

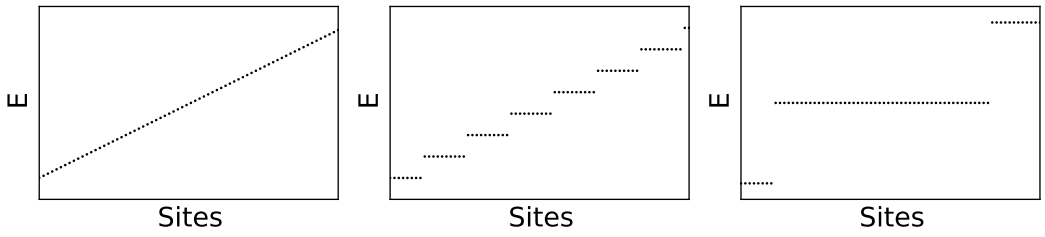
Figure 2.8 *Niche breadth σ and rank abundance distributions for the homogeneous and uniform heterogeneous environment, taking data from the final 3000 time steps.*

We see that in the homogeneous case, one single extreme specialist species dominates the environment, while in the heterogeneous case, many species coexist together, albeit with an emphasis on generalists.

Results: variable σ in simple heterogeneous and homogeneous environments

In order to visualise the resultant communities that arise from these homogeneous and heterogeneous environments, we can plot σ distribution and rank abundance distribution (Figure 2.8) plots. The σ distribution shows the number of individuals of a particular σ value, and does not differentiate between species. We will now also plot information about σ on the rank abundance plot. The average σ of the species at each rank is found for a number of replicates and colour-coded according to its σ category. Note that this averaging process may result in σ categories with low abundance not being represented on this plot.

Examining the σ distribution of the homogeneous environment (Figure 2.8a), we see that the community is made up almost entirely of extreme specialists.



(a) *g=1* **(b)** *g=10* **(c)** *g=50*

Figure 2.9 *Illustration of a section of the resource spectrum using different g (grain size).*

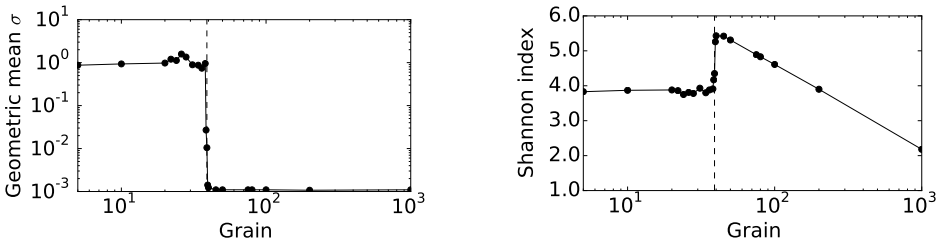
Dots represent sites. A grain refers to a number of sites with the same E value. We can see that the larger a grain is, the larger a gap in E there is between adjacent grains in E space.

Examining the rank abundance distribution of this community (Figure 2.8c), we see that this community is made up almost entirely of a single, very successful, extreme specialist species.

By contrast, the uniformly heterogeneous environment is more diverse, although dominated by generalists. The σ distribution (Figure 2.8b) shows that the vast majority of the community is generalists, with small numbers of specialist species. Examining the rank abundance distribution (Figure 2.8d), we see that high abundance species tend to be generalists, with some mild specialist species occurring at lower abundances. However, unlike in the homogeneous environment, many generalist species coexist rather than a single species dominating.

2.4.1 Environments with varying grain size

As we observe that generalists win in a heterogeneous environment, and specialists win in a homogeneous environment, it is interesting to explore what happens in environments which are intermediate between the two. We first attempt to do this by separating the environment into 'grains' of equal size. Each grain represents a small homogeneous part of the resource spectrum, i.e., a number of sites with equal E values. When the grains are very large (of the order of N, we expect to see specialists winning, since the environment is close to the homogeneous case. When grains are very small (of the order of 1), we expect to see generalists winning since the environment is close to the uniform heterogeneous case. However, it is unclear what will happen at intermediate values of grain size.



(a) *Geometric mean σ* **(b)** *Shannon diversity*

Figure 2.10 *Geometric mean σ and Shannon diversity for a range of grain sizes, using data taken from the final 10% of the simulation*

We see that a drop in geometric mean σ with increasing grain size occurs with a corresponding spike in Shannon diversity. This is because, with a small enough grain size, more specialists can coexist than generalists can as they are not directly competing. However, as grain size increases, the number of different grains correspondingly decreases, resulting in fewer specialist species coexisting and thus a lower Shannon diversity.

We define the grain size g as the size of a grain. Our heterogeneous environment is therefore of grain size g=1 (each site has a different E value). Our homogeneous environment is of grain size g=N (all sites have the same value). Figure 2.9 shows an illustration of a small part of the resource spectrum for g=1, 10 and 50. Note that as the grain size is increased, the gaps between each grain must also necessarily be increased as E varies between 0 and 1.

Results: environments with varying grain size

Simulations were run for 30,000 time steps for a range of different g. We examine the state of the community by two measures - the geometric mean σ and Shannon diversity (Equation 1.1). The geometric mean is given by Equation 2.3, where n is the total number of individuals and a_i is the σ value of individual i. The arithmetic mean is not suitable as σ varies on a geometric scale; using the arithmetic mean would cause generalists to be overrepresented.

$$\text{Geometric mean } = \sqrt[n]{\prod_{i}^{n} a_i} \tag{2.3}$$

Figure 2.10a shows the resultant geometric mean σ for a range of grain sizes. Each dot shows results of an individual simulation. When grains are small,

the community is predominantly generalist, and maintains a constant value of geometric mean σ and Shannon diversity (Figure 2.10b) across all grain sizes. At g=39, shown on both plots with a dotted line, a sharp transition is observed. The resultant community becomes entirely extreme specialist. This is accompanied by an upward spike in diversity. As grain sizes increase further, the community remains composed of extreme specialists, and the diversity decreases.

The result seen in Figure 2.10b is unsurprising if the definition of Shannon diversity is considered analytically. Taking once again the definition of Shannon diversity from Equation 1.1, we can see that here the Shannon diversity simplifies to the following, where k is the number of grains:

$$\text{Shannon diversity} = -\sum_{i=1}^{k} \frac{g}{N} \ln\left(\frac{g}{N}\right) \tag{2.4}$$

As k is simply equal to $\frac{N}{g}$, this further simplifies to:

$$\begin{aligned}
\text{Shannon diversity} &= -\sum_{i=1}^{k} \frac{1}{k} \ln\left(\frac{1}{k}\right) && (2.5)\\
&= -\sum_{i=1}^{k} \frac{1}{k} (\ln(1) - \ln(k)) && (2.6)\\
&= -\sum_{i=1}^{k} -\frac{\ln(k)}{k} && (2.7)\\
&= \ln(k) && (2.8)\\
&= \ln\left(\frac{N}{g}\right) && (2.9)\\
&= \ln(N) - \ln(g) && (2.10)\\
& && (2.11)
\end{aligned}$$

As $\ln(N)$ is a constant, we find that the Shannon diversity is direct function of $\ln(g)$, recovering what is seen in Figure 2.10b.

Our interpretation of this result is that one can view the coarse-grained environment as a series of niches, within each of which a different specialist species dominates, due to competitive exclusion [36, 72, 84, 126]. As the environment

becomes more fine-grained, the grain size decreases, and adjacent grains become more similar to each other. This means that specialists which colonise a single grain have a smaller and smaller niche on which to subsist. Once the grains are sufficiently small, and sufficiently similar to each other, specialists can be outcompeted by generalists that can utilise the resources in multiple grains. That is, despite the fact that they are less adapted to any one grain than a specialist, a generalist can build up a sufficiently large population across multiple grains to squeeze out the specialists by sheer force of number. This finding echoes previous suggestions that a minimal spatial habitat size is needed to support specialists [55]; a similar effect was also seen in the more minimal model of [115].

We also observe that the species diversity, as measured by the Shannon index, is negatively correlated with grain size for specialist-dominated communities ($g > 39$ in Figure 2.10a). This fits with the 'one specialist per grain' picture of the coarse-grained environment, since as the grain size increases, the number of niches decreases, and consequently the number of specialist species that can be supported by the environment decreases. It is interesting to note that in the fine-grained, generalist-dominated regime ($g < 39$ in Figure 2.10) neither the mean niche breath nor the species diversity change with the grain size. This reflects the fact that the dominant generalist species are by definition blind to the structure of the environment. Indeed, investigating in more detail, we find that the community composition, as defined by the distribution of μ and σ values, varies only weakly with the grain size in this regime. In this regime, one might expect the community as a whole to exhibit the characteristics of a neutral model [71]. We note that the Shannon diversity at its peak is slightly lower than the Shannon diversity for a heterogeneous environment populated only by specialists of $\sigma = 10^{-3}$. We saw that in that environment, each species had an abundance of roughly 25 (Figure 2.5a). Here, each grain is 39 sites large, which means that a specialist must have a larger habitat in order to survive when it is competing with generalists.

If we examine the time series of the community composition at g=39, we observe a dynamical transition (Figure 2.11). Initially, the community is generalist, and exists in this state for around 7000 time steps. The simulation is biased towards generalists at the beginning of the simulation, because there is no need for specific generalists to arrive via immigration, while any successful specialist must have a μ value tuned to the resources available. The generalist population then rapidly crashes, which is accompanied by first an increase in mild specialists, which then

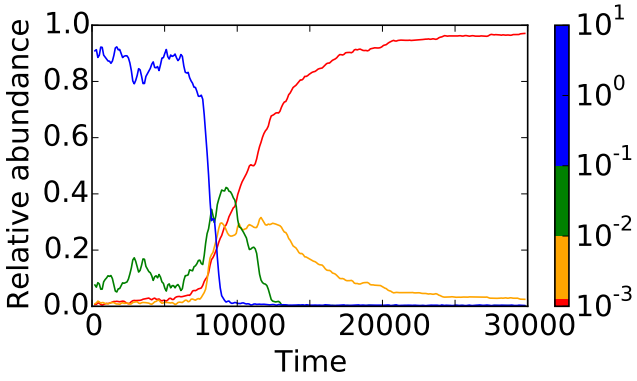


Figure 2.11 *Time series of niche breadth σ populations in a uniformly heterogeneous environment with grain size g=39*

A dynamic transition occurs. The system starts initially dominated by generalists, which coexist in a semi-stable state. However, at around 6000 timesteps, the generalist population suddenly crashes and is replaced by species that are increasingly specialist over time.

decreases, an increase in intermediate specialists, which then decreases, and finally an increase in extreme specialists, which come to dominate the community.

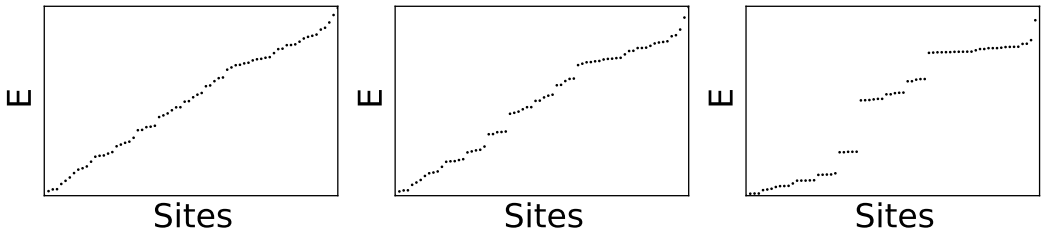
2.4.2 Environments with a discontinuously varying resource spectrum

The models for the resource spectrum which we have investigated so far are artificial in that they involve discrete grains, within which the environmental parameter E is constant. In reality, resource spectra are unlikely to be so neatly partitioned. A key question then is the extent to which insights drawn from these simple models apply to more complex resource spectra. To address this, we introduce a model of the resource spectrum in which the environmental parameter E varies with stochastic discontinuities. This model will hereby be referred to as the 'discontinuously varying resource spectrum model'.

In this model, we set the environmental value of site i (for $i = 1 \ldots N$) to

$$E_i = \frac{\sum_{j=1}^{i} x_j}{\sum_{j=1}^{N} x_j} \,. \tag{2.12}$$

where the x_i values, which are proportional to the gaps in the resource spectrum



(a) *ρ=0.01* **(b)** *ρ=0.3* **(c)** *ρ=0.4*

Figure 2.12 *Illustration of a section of the resource spectrum using different ρ.*

We can see that for a larger ρ, sites adjacent in E space are brought either closer together or pushed further apart, in a way similar to increasing g resulting in larger grain sizes but larger gaps between grains.

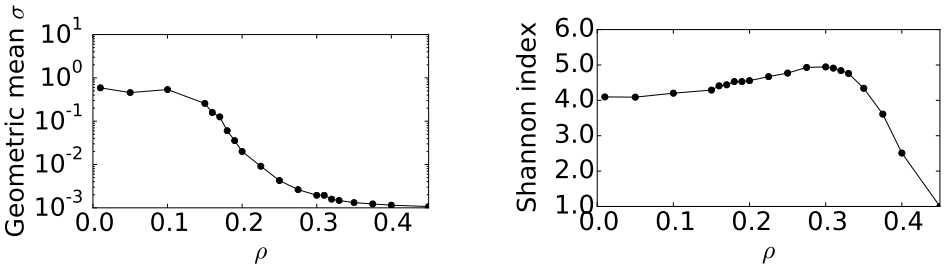
between adjacent sites) are chosen from the probability distribution

$$p(x) = \frac{1/\rho - 1}{x^{1/\rho}} \tag{2.13}$$

The resulting resource spectra are illustrated in Figure 2.12. As Figure 2.12 shows, larger values of ρ (parameter controlling the degree to which discontinuities appear in the environmental resource spectrum) correspond to more coarse-grained environments, where in this case we associate a region of slowly increasing environmental parameter, separated by large gaps as a grain. For low values of ρ, the environment is fine-grained. The power-law distribution, Eq. (2.13), has the feature that its variance is undefined when $\rho > \frac{1}{3}$; in this regime, very large gaps appear in the resource spectrum.

Results: environments with complex discontinuously varying grain size

Figure 2.13 shows the geometric mean σ of communities with varying values of ρ. The models with step-like resource spectra show a shift from generalism to specialism when moving from fine-grained to coarse-grained parameters. Consequently, in this discontinuously varying resource spectrum model, we expect to see the same shift as ρ is increased. Comparing Figure 2.13 with Figure 2.10 we indeed see consistent trends: specialism is favoured when the resource spectrum is 'grainy' (e.g. Figure 2.12c; ie spectrum contains large 'grains' with large gaps between them), while generalism is favoured for smoother resource spectra (e.g. Figure 2.12a). The shift is however more gradual than in the explicitly



(a) *Geometric mean σ* **(b)** *Shannon diversity*

Figure 2.13 *Geometric mean σ and Shannon diversity for a range of ρ in the discontinuously varying resource spectrum model*

The transition between the specialist and generalist regimes in the discontinuously varying resource spectrum model. In (a), we observe a gradual transition between from a generalist community (at small ρ) to a specialist community (at large ρ). In (b), we observe a peak in diversity for intermediate values of ρ, analogous to that seen for the uniformly grained model in Figure 2.10b.

grained model. Similarly, the diversity (as measured by the Shannon index) varies smoothly with ρ, although vestiges of the peak in diversity evident in the uniformly grained case (Figure 2.13b) remain in this case with a discontinuously varying resource spectrum. In particular, diversity initially increases as the spectrum becomes more fine-grained, before decreasing once generalists begin to dominate. Therefore, the intuitions drawn from uniformly grained, step-like models do extend to more complex resource spectra.

We can also examine the σ distributions (Figure 2.14) and rank abundance distributions (Figure 2.15) for a range of ρ values. At low values of ρ, both the σ distribution (Figure 2.14a) and rank abundance distribution (Figure 2.15a) are close to those seen when using a uniformly heterogeneous resource spectrum (as in Figures 2.8b and 2.8d), although with more specialists surviving as the resource spectrum still has some discontinuities. As ρ increases, more specialists survive, and the σ distribution shifts from generalist-dominated to specialist dominated. At ρ=0.2 (Figure 2.14b), there is wide coexistence of generalists and specialists. Examining the rank abundance distributions, we see that when there is coexistence of specialists and generalists, the most specialised species usually have an abundance that places them in the middle of the rank abundance curve, with particularly high and low abundance species being less specialised.

To establish whether this coexistence is stable, we can examine time series trajectories of species with different σ categories over time (Figure 2.16). We see

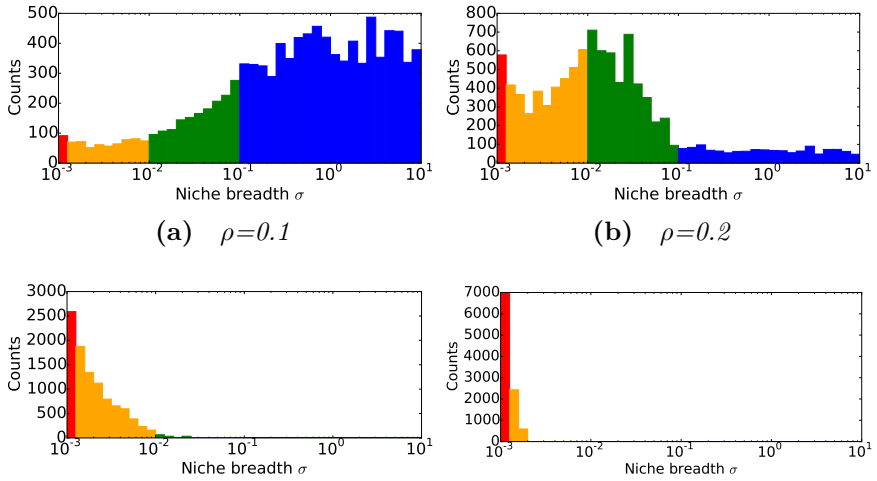


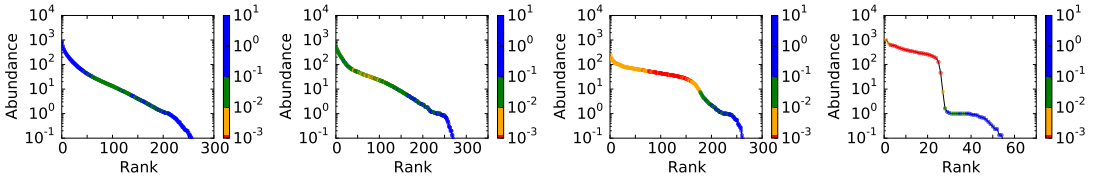
Figure 2.14 *σ distribution at a range of ρ, using data from the final 10% of the simulation*

When ρ is low (a), the σ distribution is similar to that of the case of the uniformly heterogeneous environment (Figure 2.8b). However, as ρ is increased, the distribution becomes skewed towards specialist species.

that mild and intermediate specialists can maintain a high population without being displaced by extreme specialists when ρ is sufficiently low (Figures 2.16a and 2.16b), although when ρ becomes large (Figures 2.16c and 2.16d), mild and intermediate specialists are rapidly and gradually displaced, respectively. Long-time simulations of ρ=0.2 and 0.3 (Figure 2.17) support this conclusion.

2.5 Effect of coarse-graining on specialist-generalist balance and diversity

Our results show that, in both the uniformly grained model and in a more complex discontinuously varying resource spectrum model, resource spectrum coarse-graining causes a shift from a generalist-dominated state to one that is dominated by specialist species. In the simple model, in which all grains are equal in size and in their degree of isolation, this transition is very sharp; it is preserved in the more realistic discontinuously varying resource spectrum models, but becomes broader. This observation echoes the established view in the ecology literature that large homogeneous 'patches' of habitat or of a particular resource tend to be occupied by a single species which is highly specialised to



(a) ρ=0.1 **(b)** ρ=0.2 **(c)** ρ=0.3 **(d)** ρ=0.4

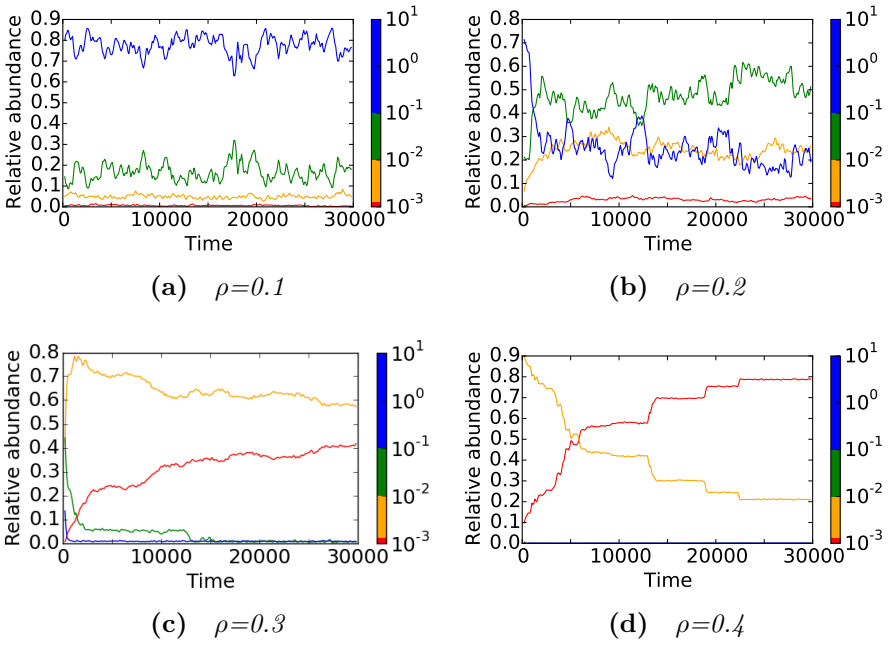
Figure 2.15 *Rank abundance distributions at a range of ρ, using data from the final 10% of the simulation.*

Rank abundance points are colour coded according to the mean σ of species at that rank abundance point. We see a general trend that more specialist species occur in the middle of the rank abundance plot, meaning that the most abundant and least abundant species tend to be more generalist.

that resource or habitat [9, 49, 59, 111]. Our results also complement previous model predictions that spatially coarse-grained habitats should favour specialists [26, 115]; here we show that the same phenomenon also occurs in a well-mixed system where it is the resource spectrum, rather than the spatial habitat, which is coarse-grained.

Additionally, our work shows that for specialist-dominated communities, an increase in resource spectrum coarse-graining is associated with a decrease in diversity. From the perspective of spatial habitat structure, this supports the view that habitat heterogeneity results in more diversity [94]. More recently, however, it has also been suggested that habitat 'microfragmentation' may result in a unimodal distribution, where diversity peaks at intermediate heterogeneity, and is lower for either very low or very high levels of heterogeneity [83, 88]. This is because very high levels of heterogeneity result in smaller habitat areas, which are not large enough to support individual specialist species, so that instead a single generalist may take over this habitat. Indeed, we see this effect in our uniformly grained and discontinuously varying resource spectrum models. Taken together, our results over the entire range of resource spectra do support the hypothesis of a peak in diversity at intermediate levels of intermediate coarse-graining.

Interestingly, in the generalist-dominated limit, for the fine-grained resource spectra, our system shows characteristics of a neutral model, in that its composition is essentially independent of the environment. In contrast, for coarse-grained resource spectra, our model is highly niche-dominated, with individual niches being occupied by specialists which exclude other species by competitive exclusion. Thus, our models provide a convenient way to explore the crossover



(a) *ρ=0.1* **(b)** *ρ=0.2*

(c) *ρ=0.3* **(d)** *ρ=0.4*

Figure 2.16 *Time series at a range of ρ*

Species of different σ values may coexist at lower ρ (a,b) but for larger ρ (c,d), the population becomes ever more specialist.

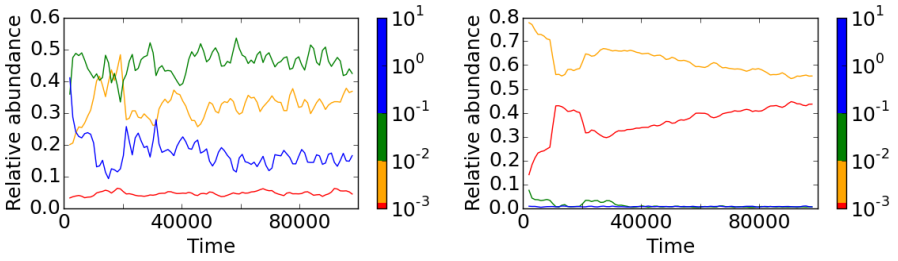
between niche theory and neutral theory; a topic which has been of increasing interest to theoretical ecologists in recent years [15, 64, 83, 117, 142, 156].

2.6 Conclusions

In this chapter, we have introduced the mechanisms of our model and established two methods of creating environments with varying amounts of graininess, one in which the grains are completely homogeneous and equally sized and spaced, another in which graininess is achieved by making a continuous resource spectrum more or less discontinuous.

In environments with a small fixed σ, corresponding to a community of specialists, we have shown that our model recreates the phenomenon of niche partitioning, core to ecological niche theory.

When we allow σ to vary, simulating competition between specialists and generalists, we have shown that environments with large discontinuities (and hence large homogeneous grains) in the resource spectrum result in specialists,



(a) ρ=0.2 **(b)** ρ=0.3

Figure 2.17 *Long-time simulations, taking a 2000 point moving average.*

We see that species continue to coexist for a long time at ρ=0.2. For ρ=0.3, the system appears to move very slowly towards an ever larger extreme specialist population.

while environments with small discontinuities result in generalists, which is in agreement with the widely accepted ecology literature.

We have used a well-mixed agent-based model of community dynamics, with a continuous range of specialist-generalist strategies, to investigate the interplay between the environmental resource spectrum and community composition. Even in this well-mixed model, coexistence between specialist and generalist species is possible because different specialist species can occupy different parts of the resource spectrum. Our results should be most relevant to natural environments where dispersal rates are high (e.g. birds foraging in a heterogeneous forest or aquatic environments with high mixing rates) or where a spectrum of resources is available in a single location (e.g. the light spectrum, or food particles of a range of sizes or types). We find that resource spectrum coarse-graining favours specialists over generalists, since specialists require niches of a critical size in order to support a stable population.

In the following chapter, we will continue to use this model, using the two environments we have established, and introduce a third environmental model to investigate a new proposed mechanism of community assembly.

Chapter 3

Resource Spectrum Engineering

In the previous chapter we introduced our well-mixed model, and outlined two types of environment in which the graininess of the resource spectrum can be controlled. We found that an environment which is more coarse-grained favours specialists, and an environment which is more fine-grained favours generalists.

In the case of the uniformly grained model, we saw that the system can exist in a generalist-dominated state for a long time before undergoing a sudden dynamical transition into a specialist-dominated state. In this chapter, we uncover the reasons why this transition occurs.

First, we extend the uniformly grained model such that grain size and the gaps between grains can be varied independently, to establish the effect of increasing both grain and gap size on triggering the transition to specialism. Second, we investigate if this effect is present in our discontinuously varying resource spectrum model, by removing highly specialised species from the simulation and examining the effect on the resultant community composition. Taking these together, we conclude that the presence of highly specialised specialists has an advantageous effect on the abundance of mild specialists through a new mechanism which we term 'resource spectrum engineering'. In this mechanism, highly specialised species occupy attractive sections of the resource spectrum and cannot be displaced. These sections of the resource spectrum are then effectively off-limits to less specialised species, resulting in a perceived resource spectrum which is more discontinuous, which advantages mild specialists.

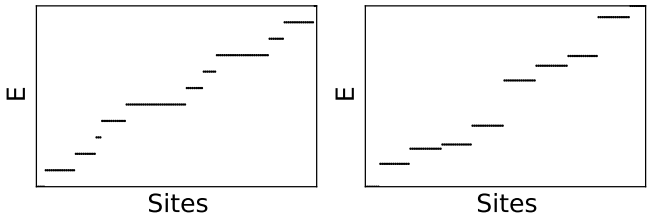
3.1 Variable grain size and variable gap size models

In our uniformly grained model of Chapter 2, a dynamical transition from generalism to specialism is triggered when $g = 39$ (Figure 2.10a). Examining a dynamical trajectory for this grain size reveals that the system initially exists in the generalist-dominated state, until a transition to a specialist-dominated state occurs (Figure 2.11). In this transition, the mild specialist population first peaks, followed by the intermediate specialist population, followed by the extreme specialist population which eventually comes to dominate the community. At lower values of g ($g < 39$), the transition never occurs, and the generalist-dominated community prevails for the duration of the simulation. At higher values of g ($g > 39$), the transition to a specialist-dominated community happens immediately at the start of the simulation. Only at g values very close to 39 can the dynamical transition be observed.

The reasons for this transition, and why it occurs via this sequence of organism types, are unclear. Clearly an increase in grain size results in the transition being more likely, but as grain size is inherently linked with gap size in this model (as N sites must be distributed in some way between 0 and 1), we do not know if it is solely in the increase in grain size, or the increase in gaps between grains, that makes the transition occur.

Another question is what event actually causes the transition. A naive explanation for extreme specialists displacing generalists could be that suitable extreme specialists enter the simulation through immigration, each colonising a suitable grain as they arrive. However, this would result in a linear increase in extreme specialist numbers, rather than the dramatic system-wide transition that we observe. This scenario also would not explain the peaking of mild and intermediate specialist populations that we see during the transition.

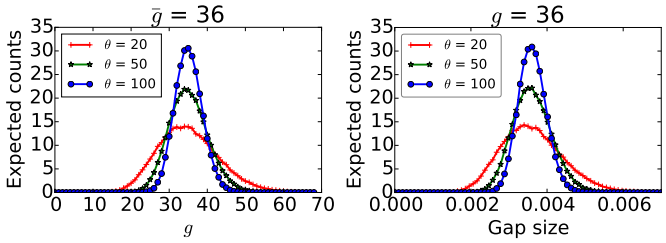
It is generally accepted that large grains result in specialists [9, 49, 59, 111], but the role of gaps in a resource spectrum has been less well studied. To distinguish the roles of grain size versus gap size, in this section we introduce a model of the environmental resource spectrum that allows grain size to vary while gap size remains fixed, and vice versa, so as to isolate the effects of these two factors on the community composition. We then examine the populations that occupy particularly large grains or grains adjacent to large gaps and from this deduce



(a) *Variable grain size* **(b)** *Variable gap size*

Figure 3.1 *Illustration of variable grain size and variable gap size resource spectra*

A simplified illustration of environments with variable gap and grain sizes. In (a), grain size varies while gap size is held constant. In (b), the reverse is true.



(a) *Variable grain size* **(b)** *Variable gap size*

Figure 3.2 *Distribution of expected grain and gap sizes and $\bar{g}$ and g, respectively.*

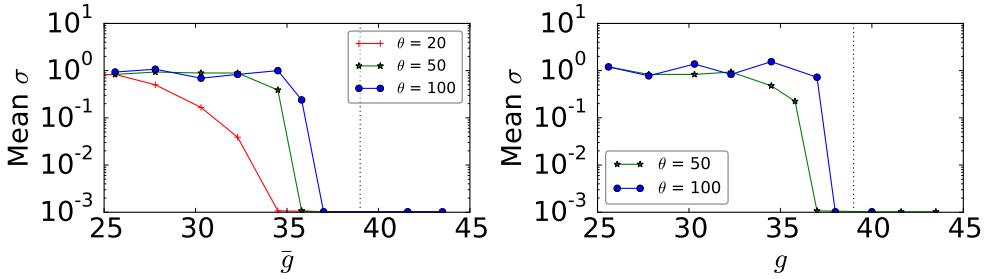
When θ is high (blue line), the expected distribution of grain and gap sizes is narrow, and when it is low (red line), the expected distribution is wider, resulting in more extreme values. As θ approaches ∞, the resultant environment approaches the uniformly heterogeneous environment.

what causes the transition in our original uniformly grained model.

3.1.1 Assignment of environmental values

In this model, we fix the total number of grains, k, and *either* the grain size $g = N/k$ *or* the size of the gaps between grains $1/k$ (Figure 3.1). In the former case we draw k gap sizes from the gamma distribution

$$p(x) = \frac{x^{\theta-1}e^{-x}}{\Gamma(\theta)} \quad \text{where} \quad \Gamma(a) = \int_0^\infty t^{a-1}e^{-t}\mathrm{d}t \tag{3.1}$$



(a) *Variable grain model* **(b)** *Variable gap model*

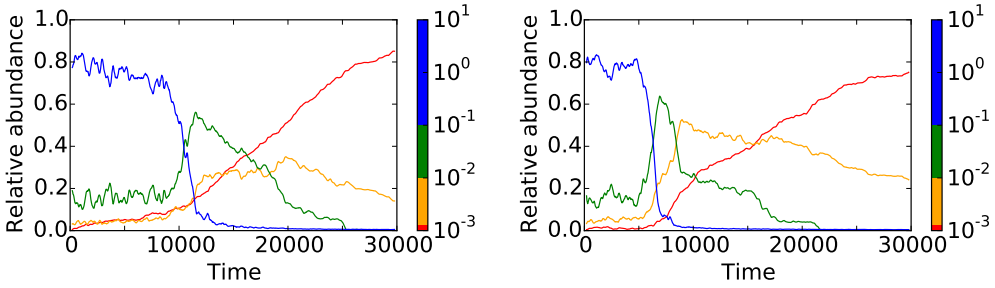
Figure 3.3 *Phase transition in variable grain and variable gap size models*

The transition between the specialist and generalist regimes in cases of variable grain and gap sizes for different values of the variability parameter θ (see main text). The transition point of $g = 39$ from the uniformly grained model is shown by the black dotted line.

and normalise them so that the sum over all k gap sizes is equal to 1. In the latter case, we draw k grain sizes from the same gamma distribution, and normalise so that their sum equals N, the number of sites and they are rounded to the nearest integer. The gamma distribution was chosen because it has a single parameter θ that controls the variability of the gap or grain size. The key feature of this distribution is that the ratio of the standard deviation to the mean is $1/\sqrt{\theta}$. Hence, when θ is large, there is only limited variation in grain or gap size; when θ is small we expect to see both large and small gaps or grains in the resource spectrum. Note that when discussing the variable grain size model, we will refer to the mean grain size $\bar{g}$, for ease of comparison with the other models. Figure 3.1 shows a portion of the resource spectrum for both the variable gap and grain size models. Figure 3.2 shows the distribution of grain sizes or gap sizes for those models with $\bar{g}$ and $g = 36$, respectively, for a range of values of the variability parameter θ.

3.1.2 Variability in grain or gap size favours transition to specialism

The models with variable grain size and variable gap size show the same transition from generalist-dominated to specialist-dominated communities as the simpler model of the previous chapter. Figure 3.3 shows the geometric mean niche breadth at long times (30 000 generations) as a function of grain size $\bar{g}$ or g, for simulations with $N = 10000$ sites and several values of the variability parameter θ. A low



(a) *Variable grain model* **(b)** *Variable gap model*

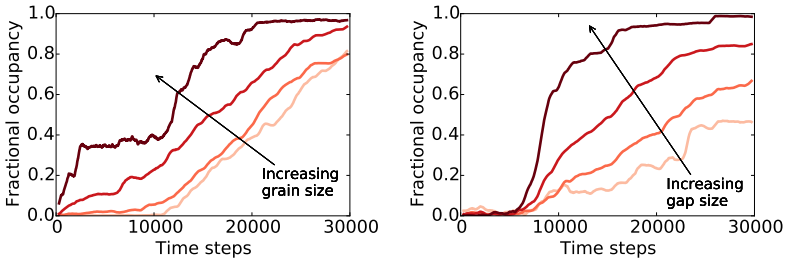
Figure 3.4 *Time-series showing the dynamical transition from generalism to specialism in both variable-grain and variable-gap models.*

The time series shows the relative proportion of species belonging to different σ categories, as seen in the key to the right of each figure. In both cases, a clear transition point can be seen in which the population of generalists rapidly drops, and the population of all types of specialist start to increase.

value of θ means more variability. More variability in the environment results in the transition to specialism occurring at lower g or $\bar{g}$: in other words, more variability results in specialists becoming established more easily.

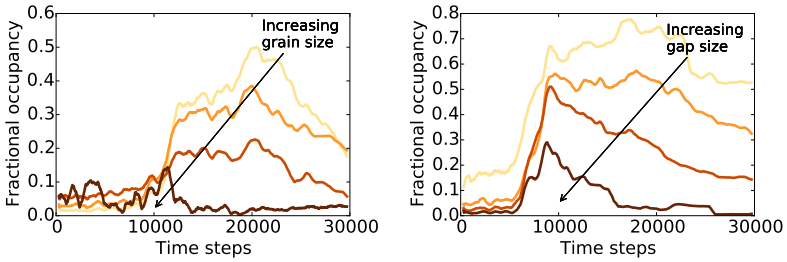
Why do more variable resource spectra favour specialists? In our simulations of the idealised model in the previous chapter, we saw that when the grain size reached the threshold of $g = 39$, extreme specialists colonised every grain. This seems to suggest that these specialists need a grain size $g > 39$ on which to thrive. However, in the variable grain model, we see that the entire environment becomes populated by extreme specialists at a smaller mean grain size, $\bar{g} = 36$ (Figure 3.3a). Moreover due to the variability in grain size, many grains which become populated by extreme specialists must be much smaller than this (Figure 3.2a). Thus, grain size variability appears to result in more complex behaviour, where the occupation of larger grains by specialists makes smaller grains more attractive to other specialists.

To shed light on this situation, we plot in Figure 3.4 dynamical trajectories for our simulations for the parameter set $\bar{g}$ or $g = 36$, θ=40, which results in a system very near to the transition point. In these trajectories the system starts off in a generalist regime and transitions into a specialist regime. By examining how this transition happens dynamically, we seek to find the causes of the phase transition as a whole.



(a) *Variable grain size, extreme specialists*

(b) *Variable gap size, extreme specialists*



(c) *Variable grain size, intermediate specialists*

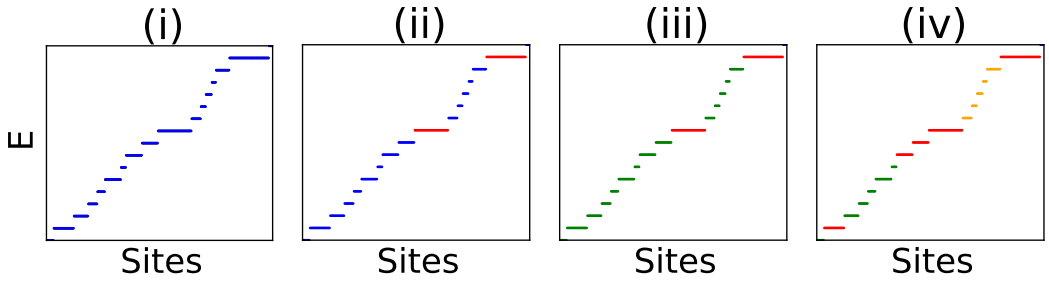
(d) *Variable gap size, intermediate specialists*

Figure 3.5 *Occupation of different grain and gap sizes by extreme and intermediate specialists*

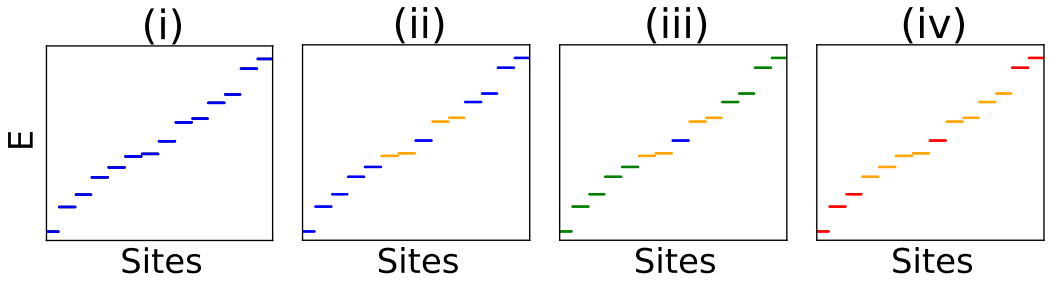
Dynamical trajectories showing the fractional occupation of different types of grain by extreme ((a),(b)) and intermediate ((c),(d)) specialists for simulations with g or $\bar{g} = 36$ and $\theta = 40$. For the variable grain size model (a) and (c), grains are sorted into different categories depending on their size. These categories are determined by taking equal quartiles across the whole range of grain sizes present. For the variable gap size model (b) and (d), gaps are sorted into size categories by the same method, and the adjacent grain to each gap placed into that size category. As each grain has two adjacent gaps, each grain is counted twice.

3.1.3 Transition dynamics for the variable grain model

Figure 3.4 shows transition trajectories in which the system changes from generalist dominated to specialist-dominated ($\bar{g} = 36$, θ=40). The parameters of $\bar{g}$ or $g = 36$ and $\theta = 40$ have been chosen to show behaviours very near to the transition point (Figure 2.10). σ categories are shown at the right hand side according to the colour scheme outlined in Figure 2.7. Generalists initially dominate, until a time is reached where the generalist population collapses. From then on, increasingly specialist populations come to dominate in a successive fashion. In both cases, extreme specialists are over-represented



(a) *Variable grain model*



(b) *Variable gap model*

Figure 3.6 *The proposed mechanism behind the transition in the variable grain size and variable gap size models. (i), (ii), (iii) and (iv) show progression over time.*

In both cases, initially, the population is dominated by generalists. Attractive grains are then filled by opportunistic specialists, which causes gaps in the resource spectrum. These gaps can then be exploited by mild specialists, which temporarily dominate, before being outcompeted by more specialised species.

among the specialists pre-transition, and come to almost completely dominate post-transition with their population ever-increasing, despite representing only 2.5% of the strategies available.

We now examine in more detail the grains on which the extreme specialists reside, for the variable grain size model. In Figure 3.5a, grains are categorised by size and the fractional occupancy of extreme specialists is shown for each category of grain size. Pre-transition, only large grains are inhabited by extreme specialists. Post-transition, extreme specialists also take over smaller grains.

Based on these results, we hypothesise the following: Initially, almost all grains are occupied by generalists (Figure 3.6a(i)), due to the starting conditions of the simulation. Over time, large grains become occupied by extreme specialists (Figure 3.6a(ii)) which arrive by immigration. As it is difficult to compete with an extreme specialist living on a large grain of its preferred habitat, this creates

gaps in the resource spectrum which is available for other species. Because generalists need to occupy multiple grains in order to compete, gaps in the resource spectrum disfavour generalists, which provides an opportunity for mild specialists to subsist on sets of similar grains, which are separated at each end by grains occupied by extreme specialists (Figure 3.6a(iii)). This results in a spike in the population of mild specialists as we see in Figure 3.4a. Once the mild specialists have triumphed over the generalists, competition pressure is reduced for all of the specialists and the environment fills with extreme specialists on the largest grains and intermediate specialists on the remaining small grains ((Figure 3.6a(iv)). With time, extreme specialists encroach on the intermediate specialists, eventually resulting in domination by extreme specialists.

3.1.4 Transition dynamics for the variable gap size model

We now perform the same analysis for the variable gap size model. Once again we examine transition dynamics for a simulation condition on the cusp of transition (g = 36, θ=40), Figure 3.4b. Our simulation begins in the generalist regime, before the dynamical transition occurs which results in a specialist-dominated community.

As before, we wish to examine where extreme specialists reside during the transition. Because it is gap size which varies in this model, we divide the gaps in the resource spectrum into categories according to their size, and assess which species are found on grains which are adjacent to gaps of a given size. We note that each individual is associated with two gaps, one either side of the grain on which it sits, so that each individual is counted twice in this analysis (Figure 3.5b).

Figure 3.5 shows that, prior to the transition, occupancy of grains by extreme specialists does not depend strongly on the size of the adjacent gaps; few of the grains are occupied by the extreme specialists regardless of the size of their adjacent gaps. Post-transition, however, extreme specialists do preferentially occupy isolated grains, which are adjacent to large gaps. Interestingly, if we instead track the intermediate specialist population (Figure 3.5d), we see that both prior to the transition and post-transition, intermediate specialists tend to occupy grains adjacent to small gaps (i.e. grains that are less isolated in the resource spectrum).

Our results seem to suggest that similar, but not identical, dynamics occur in the variable gap size model as those found in the variable grain size model. Figure 3.6 shows a proposed mechanism for the transition in both cases. For the variable gap size model, initially, all grains are occupied by generalists (Figure 3.6b(i)), due to the initial conditions of the simulation. Pairs of grains that are separated only by very small gaps are then filled by intermediate specialists, which can use them as a single niche (Figure 3.6b(ii)). This creates gaps in the resource spectrum that allow mild specialists to displace the generalists (Figure 3.6b(iii)). Once the system is in the specialist regime, intermediate and extreme specialists can encroach on the mild specialists at the edges of their domains where competition pressure is less strong, eventually displacing them entirely (Figure 3.6b(iv)).

The fact that extreme specialists preferentially occupy isolated grains post-transition (i.e. when the system is specialist-dominated), but not pre-transition (when it is generalist-dominated; Figure 3.5b), is also interesting. This suggests that specialists gain a definite advantage from occupying niches that are isolated in the resource spectrum, but only if their competitors are other specialists. This may be because competition pressure from generalist competitors is unaffected by the degree of isolation of a particular niche, since a generalist's niche breadth spans the entire environment.

3.1.5 Resource spectrum engineering by extreme and intermediate specialists

Taken together, our results for the step-like resource spectrum models lead to the following hypothesis:

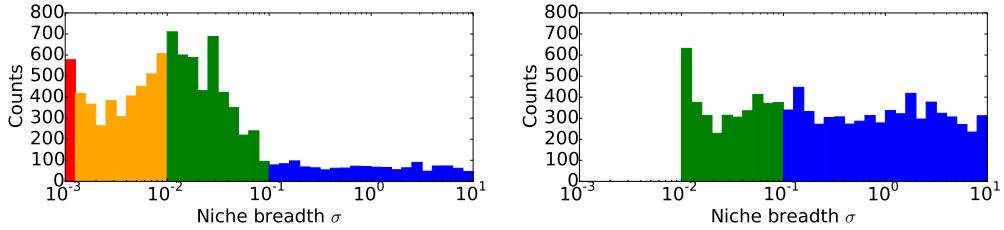
Heterogeneity in the resource spectrum produces niches which are able to sustain a specialist population; there are either large grains that can support extreme specialists (variable grain size model) or grains separated by small gaps that can support intermediate specialists occupying several adjacent grains (variable gap size model). Establishment of specialist species on these niches creates gaps in the resource spectrum available to other species, since these occupied parts of the resource spectrum are 'off-limits' to less specialised species. This effectively partitions the remaining available resource spectrum, providing mild specialists with isolated habitats that are protected from competition pressure from generalists, since generalists reply for competitiveness on being

able to occupy patches consisting of multiple grains. Once competition pressure from generalists has been eliminated, competition proceeds among the various specialist species. During this phase of the dynamics, mild specialists are weaker at the borders of their domains, so that intermediate and extreme specialists can encroach. With time, the mild specialists die out entirely, leaving a community which is dominated by intermediate and extreme specialists. We call this process of specialists occupying attractive parts of the resource spectrum and thus changing the competitive environment for other species, 'resource spectrum engineering'. This is analogous to 'ecosystem engineering' in which species such as beavers change the physical environment, changing the ecosystem for themselves and other species. In this case, the physical environment is unchanged, but the resources available to other species is changed.

3.2 Resource spectrum engineering in the discontinuously varying resource spectrum model

In the variable gap and grain size models, we found evidence for resource spectrum engineering by examining directly the occupancy of grains by highly specialised species (extreme and intermediate specialists, $10^{-3} < \sigma < 10^{-2}$). We now return to the discontinuously varying resource spectrum model of the previous chapter (Section 2.4.2). In this discontinuously varying resource spectrum model, there are no easily-defined grains, so this is not possible (Figure 2.12). To investigate the role of highly specialised species in a different way, we ran a new set of simulations using the same simulation parameters, but preventing the creation of any intermediate or extreme specialists, either in the initial conditions or as immigrants. A drastic change in behaviour as a result of excluding the highly specialised species would suggest that a similar resource spectrum engineering mechanism may be at play in this model.

We indeed find that when specialised species are removed from the simulation (i.e., are never created), there is a large effect on the resultant σ distribution of the species that remain. Removal of highly specialised species from the simulation appears to give the generalists an advantage over the mild specialists. Figure 3.7 shows our results for the full community composition (distribution of niche



(a) *All species*

(b) *Extreme and intermediate specialists removed*

Figure 3.7 *Distributions of σ for the discontinuously varying resource spectrum model*

Full community composition (distribution of niche breadth σ), after 30,000 generations, for the discontinuously varying resource spectrum model for ρ=0.2. (a) full range of niche breadth, (b) a niche breadth range that excludes extreme and intermediate specialists $10^{-2} < \sigma < 10^{1}$. We can see that the absence of intermediate and extreme specialists results in generalists competing more effectively against mild specialists.

breadths σ) after 30,000 generations, for simulations with the full range of niche breadths ($10^{-3} < \sigma < 10^{1}$; Figure 3.7a), compared to those where we exclude extreme and intermediate specialists (i.e. including only niche breadths $10^{-2} < \sigma < 10^{1}$; Figure 3.7b). Indeed, for this model, and for this value of the coarse-graining parameter $\rho = 0.2$, the community is dominated by mild specialists when the full range of niche breadths is included (Figure 3.7a). However, when the extreme specialists are removed (Figure 3.7b), mild specialists no longer prosper; instead the community becomes generalist-dominated. More specifically, when extreme and intermediate specialists are present, the generalist population is around 13% of the total population, much smaller than the population of mild specialists. In contrast, for simulations where the extreme and intermediate specialists are excluded, the generalist population is much higher, around 63% of the total population, and dominates over the mild specialist population. Indeed, the mild specialists, as a fraction of the total population, drop from 43% in the presence of highly specialised species to 36% in their absence, even though the chances of picking a mild specialist as an immigrant are higher in the latter case. Thus, the presence of highly specialised species clearly favours mild specialists and disfavours generalists.

To explore the effect that specialised species have on less specialised species, we can run a range of simulations with varying lower σ bound and find the resultant change in relative abundance for mild specialists and generalists (Figure 3.8). The

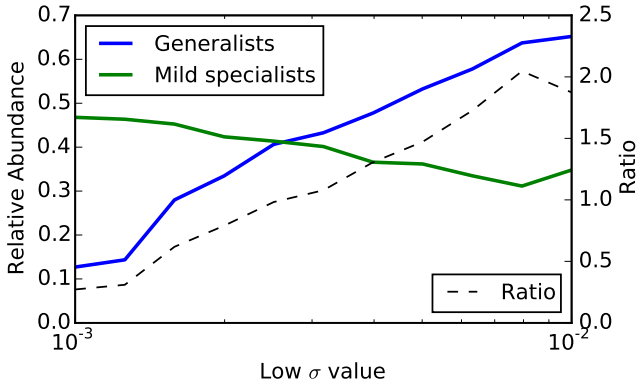


Figure 3.8 *Relative abundance of mild specialists and generalists and the ratio found at the end of simulations with a range of lower σ bounds.*

We can see that as the lower bound on σ rises (meaning less extreme and intermediate specialists in the simulation) the resultant ratio of generalists to mild specialists rises. This shows that the presence of more specialised species in the simulation aids mild specialists.

relative abundance of both mild specialists and generalists is plotted on the left hand axis, while the ratio between the two is plotted on the right hand axis (dotted line). We can see that the ratio of generalists to mild specialists increases linearly from 0.25 (one generalist to every four mild specialists) to 2 (eight generalists to every four mild specialists) as the lower bound of σ increases.

From this we can infer that the presence of extreme and intermediate specialists confers an advantage to the mild specialists. We have established that specialists are at an advantage when the resource spectrum is highly discontinuous, in both the uniformly grained and discontinuously varying resource spectrum models.

We now run a set of simulations for a range of ρ in which only mild specialists and generalists are included ($10^{-2} < \sigma < 10^{1}$). Figure 3.9 shows a comparison between the transition between generalism to specialism using a full range of σ (as in Figure 2.13a) and the transition using only mild specialists and generalists. In the simulation with full σ range (dotted line), the transition begins to happen around ρ=0.15, but in the case where more specialised species are removed, the transition does not start to happen until around ρ=0.2. This suggests that mild specialists need more discontinuous resource spectra to outcompete generalists than more specialised specialists do. However, mild specialists benefit more when those more specialised species are present. We therefore speculate that the presence of highly specialised species causes the available resource spectrum to appear more discontinuous for less specialised species, giving mild specialists an

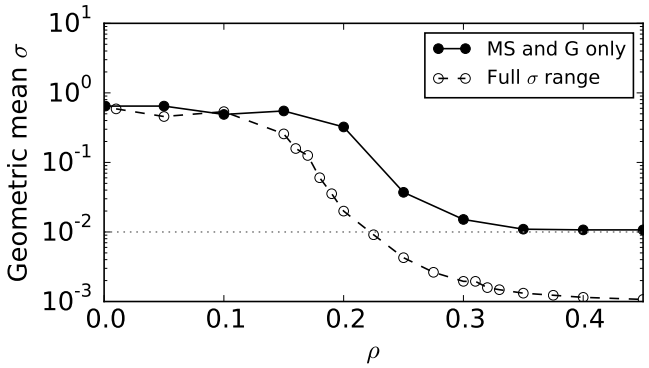


Figure 3.9 *Comparison of transition from generalism to specialism in simulations with σ limited to $10^{-2} < \sigma < 10^{1}$ (extreme and intermediate specialists removed)*

We can see that as the lower bound on σ rises (meaning less extreme and intermediate specialists in the simulation) the resultant ratio of generalists to mild specialists rises. This shows that the presence of more specialised species in the simulation aids mild specialists.

advantage over generalists and providing another example of resource spectrum engineering. More specialised species create gaps in the resource spectrum that disadvantage generalists and allow mild specialists to subsist in parts of the spectrum that lie between these gaps. Thus our results suggest that resource spectrum engineering by specialist species is not confined to simple models with step-like resource spectra, but is also a feature of this discontinuously varying resource spectrum model, with a continuous resource spectrum.

3.3 Discussion

Based on simulations in which we track community composition in time, and in which we manipulate the allowed range of specialist/generalist strategies, we postulate that the dominance of specialists in moderately coarse-grained environments can be due to a 'resource spectrum engineering' community assembly mechanism. In this mechanism, the occupation of a few niches by highly specialised species in the community can favour the occupation of other niches by specialists, by changing the 'effective resource spectrum' such that it is less favourable to generalists. Thus, the interplay between resource spectrum and community composition occurs not at the level of individual niches, but rather at the level of the entire spectrum. The presence of highly specialised species on a

few niches increases the size of the gaps between the remaining available niches, making them more isolated and thus less favourable for generalists. This picture is somewhat reminiscent of the concept of ecosystem engineers: species that modify the environment in a way that has a significant impact on other species [79]; however here the effect is not directly on an environmental parameter but rather on the distribution of the environmental parameter across the entire habitat.

Although to our knowledge this concept of resource spectrum engineering has not previously been discussed in ecological models, similar effects have been noted in the empirical literature. For example, in a study of bird communities, Julliard et al [82] found that specialised species are more abundant if the rest of the local community is specialised to the same degree.

We have shown that this phenomenon occurs in a variety of different resource spectrum models, suggesting that it may be quite general. It is also important to note that this is a phenomenon that can only be seen in models which include a continuum of strategies, from extreme specialism to generalism; we do not expect it to be evident in simpler models which include only the interplay between discrete, idealised, 'specialist species' and 'generalist species'.

3.3.1 Importance of niche isolation as well as niche size

An important concept that emerges from our work is that the gaps between grains in the resource spectrum play a key role, as well as the grain sizes. In our model with variable gap size, even though the grain sizes are identical, grains which are particularly isolated are more likely to be occupied by extreme specialists. This is because those species that specialise on a particularly unique type of resource (i.e., isolated in the resource spectrum) face less competition from intermediate specialists than those which specialise on a resource that is surrounded by similar resources. This is in agreement with the ecological literature, as it is generally accepted that specialists appear in peculiar or extreme environments [23, 154], such as those that are particularly dry or high altitude.

3.3.2 Relevance to natural communities and ecosystem management

Here, for simplicity, we have considered a static environment. Previous work shows that static environments favour specialists, while temporally fluctuating environments favour generalists [163, 173]. At present, the models that include defined, homogeneous grains bestow a perhaps unfair advantage to the extreme specialists, and small amount of more realistic temporal fluctuation may result in more of a balance between extreme and intermediate specialists in the specialist regime. However, we do not believe that small temporal fluctuations would affect our overall findings, although large fluctuations (such as a mass extinction event) may cause a shift from the specialist regime to the generalist regime, and would be a very interesting area for further study.

Our work also relies on the assumption, often made in models, that specialists and generalists can be neatly classified by their niche breadth along a single niche axis, and that a tradeoff exists between niche breadth and fecundity under optimal conditions [47, 49, 59, 84, 162, 171]. In nature, of course, the situation is more complicated: species will experience resource spectra along several niche axes simultaneously and may specialise or generalise on different ones. Indeed, measuring and comparing niche breadth between different species and environments in a quantitative way is notoriously difficult [47, 59, 84]. Moreover, difficulties also exist in quantifying the heterogeneity of an environment; one might classify the nature of an environment in terms of species requirements, i.e. its 'functional heterogeneity' [52], or in terms of natural versus managed patches and their connectivity, or in terms of their fragmentation with regard to species movements, etc [52]. Our aim in this work has been to provide basic understanding rather than direct applications to real-life ecosystems.

Nevertheless, we believe that our work does have potential implications for ecosystem management. Attempts to prevent species loss sometimes involve the creation of specific habitats aimed at supporting at-risk species [22, 108, 133]. Our work shows that this has implications not only for the specialist species that are targeted but also for the entire ecosystem, since creating such highly specialised niches may crowd out generalist species also from other, habitats that are not well suited to the specialists (the small grains in our models). This is not an effect that can easily be controlled, given that the results are diffuse, extending over the entire community.

Taken more broadly, our work highlights the fact that specialist-generalist competition does not simply happen on a local scale over specific resources but can have wide reaching implications across the entire resource spectrum of an ecosystem.

Chapter 4

The effect of spatial dispersal on the specialist-generalist balance

In this chapter, we investigate dispersal. In the previous two chapters, all species had infinite dispersal lengths, and so communities were well-mixed. Now, by introducing the life-history trait δ which describes the dispersal range of a species, we can investigate the effects of limiting juvenile dispersal such that they may only be able to establish themselves in a location that is spatially close to the parent.

As we outlined in Section 1.4.3, the relationship between long and short dispersal abilities and specialism and generalism is unclear. Specialism may be associated with long-range dispersal, in order to seek out attractive sites, or it may be associated with short-range dispersal, as attractive sites are likely to be found close together anyway. Both associations have been predicted by theoretical models. In this chapter, we will investigate this trade-off.

If dispersal of species is limited, the arrangement of environmental sites in space becomes important. First, we will use uncorrelated environments in which different types of sites are randomly placed in space. Using these environments, we will first hold a dispersal distance δ constant for all species, and find the specialist-generalist balance for each community that results for different values of fixed δ. Next, again on the uncorrelated environment, we will reverse this by holding σ constant and find the long- and short-range dispersal balance of the communities that arise.

We will then create environments with a short-range correlation by rearranging sites in space such that similar sites are found spatially close together, and repeat the same set of simulations as on the uncorrelated environment. From this, we will establish the effect of how the environment is arranged in space on the specialist-generalist balance and long- and short-range dispersal balance.

4.1 Including the dispersal limitation δ in the model

In order to limit the dispersal of our species, we will now introduce a new life-history trait which was mentioned in Section 2.1: δ.

Each species k has a δ value, which controls how competitive its juveniles are at a distance from the parent, which takes the form of a decaying exponential which we term D_k (Equation 4.1). If X is the size of the axis of the environmental lattice, (X=100 in all simulations discussed here) and z is the distance between the location of species k, which is located at (x, y) and the site onto which its offspring is dispersing (located at (i, j)) (Equation 4.2), then we define the function D_k as

$$D_k(\delta_k, z) = \exp\left(-\frac{z}{\delta_k \frac{X}{2}}\right) \tag{4.1}$$

where

$$z = \sqrt{(\Delta(x,i)^2 + \Delta(y,j)^2)} \quad \Delta(x,i) = \begin{cases} |x-i|, & \text{if } |x-i| < \frac{X}{2} \\ X - |x-i| & \text{otherwise} \end{cases} \tag{4.2}$$

Long-range dispersal comes at a cost. For example, wind-dispersed seeds may travel a longer distance only if the seed itself is less heavy, which results in a lesser chance of establishment where the seed lands [157]. It is open to debate how to model a realistic 'cost' to dispersal. We choose to assume a trade-off between dispersal and fecundity such that the integral of D_k for a species k must sum to one over the entire lattice. However, it should be noted that this is a huge simplification of the multiple factors that contribute to long distance dispersal

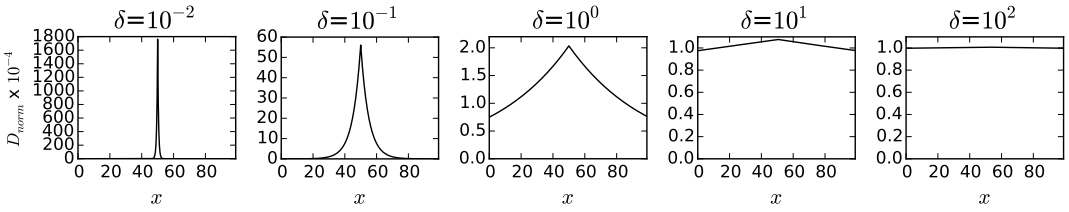


Figure 4.1 *D_{norm} over 1D for a range of δ, for an individual located at x=50.*

We see that as δ increases, the range that the juveniles of an individual may spread increases, but the probability of establishing at a site within their range decreases.

being costly in most species. Travis et al [157] argue that modeling dispersal cost in this way is too simplistic and each cause of cost to long distance dispersal should be considered individually; however, for reasons of time constraints, we have chosen to use a single dispersal kernel that encompasses all costs to long distance dispersal in one.

Our kernel results in species with a low dispersal range having a higher fecundity for offspring near to the parent than species with a higher dispersal range. The normalisation factor for a species k is dependent only on its δ value and is given by the following equation.

$$\text{Normalisation} = 4 \sum_{i=0}^{\frac{X}{2}} \sum_{j=0}^{\frac{X}{2}} \exp(-\frac{\sqrt{i^2 + j^2}}{\delta_k \frac{X}{2}}) \tag{4.3}$$

This results in a function $D_{k\text{norm}}$, which gives a measure of the probability of a juvenile of species k with δ_k establishing itself at a distance z from its parent.

$$D_{k\text{norm}} = \frac{\exp\left(-\frac{z}{\delta_k \frac{X}{2}}\right)}{\text{Normalisation}} \tag{4.4}$$

δ falls between 10^{-2}, which indicates very low-range dispersal, and 10^2, which indicates very high range dispersal. δ is selected for each new species by picking randomly from a uniform distribution between -2 and 2 and then raising 10 to the power of that number. This results in a very wide range of dispersal lengths; a species with $\delta = 10^{-2}$ may only be able to disperse to a location closer than two sites away, while a species with $\delta = 10^2$ may disperse anywhere with equal probability of establishment.

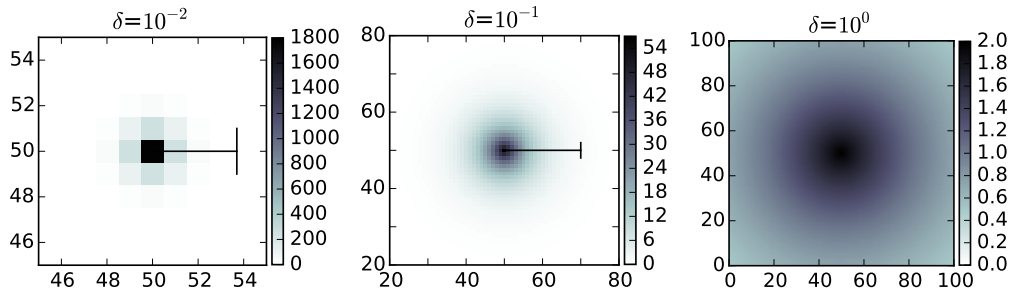


Figure 4.2 *D_{norm} over 2D for a range of δ for an individual located at x=50, y=50.*

We see that as δ increases, the range that the juveniles of an individual may spread increases, but the probability of establishing at a site within their range decreases.

Figure 4.1 shows the D_{norm} as a function of x for a species which is located at x=50, for various values of δ. When δ is small (10^{-2}), the juveniles of this species have a very high probability of establishing themselves close to the parent, but this probability falls to essentially zero at distances greater than about 2 lattice sites. As δ increases, juveniles have a greater probability of establishing themselves across the environment, but a lower probability of establishing themselves close to the parent than a species with a smaller δ. When δ is high (10^2), distance becomes irrelevant and, if all species have very large δ values, the model becomes well-mixed. Figure 4.2 shows a heat map of D_{norm} in 2D for a range of small δ, for a hypothetical species located at (x=50, y=50). For $\delta = 10^{-2}$ and $\delta = 10^{-1}$, a line extending from location of the species indicates the distance over sites that its D_{norm} is greater than 1. This will be referred to in the rest of the chapter as the 'range' of a particular species; beyond this distance, a short- or medium-range disperser will be at a disadvantage to a long-range disperser with $\delta = 10^2$.

Having established the effects of the dispersal parameter δ on the fecundity of a species, we can now combine this with the effects of the original juvenile survival function (Equation 2.2, depending on the niche breadth σ and ideal habitat μ values of a species) which was established in Chapter 2 by multiplying the two equations together. This gives an overall juvenile survival function $F(\mu, \sigma, \delta; E)$:

$$F(\mu, \sigma, \delta; E, z) = \frac{\exp\left(-\frac{\Delta(\mu,E)^2}{2\sigma^2}\right)}{\int_{-\frac{1}{2}}^{\frac{1}{2}} \mathrm{d}x \exp\left(-\frac{x^2}{2\sigma^2}\right)} \times \frac{\exp\left(-\frac{z}{\delta \frac{X}{2}}\right)}{\text{Normalisation}} \tag{4.5}$$

$$\text{where } \Delta(\mu, E) = \begin{cases} |\mu - E|, & \text{if } |\mu - E| < 0.5 \\ 1 - |\mu - E| & \text{otherwise} \end{cases}$$

The probability of establishment of a juvenile at a site with environmental value E, located a distance z away from its parent is therefore equal to its parent's full juvenile survival function $F(\mu, \sigma, \delta; E, z)$, divided by the sum of the evaluation of all other juveniles' juvenile survival functions at the same site. Technically, every parent in the environment contributes a juvenile to compete at every site, resulting in the same $N - I$ juveniles at each site competing as in the well-mixed model. However, if δ is low for all species, many juveniles will have a juvenile survival function evaluated as so close to zero as to effectively not be competing as their parent is located too far away.

Computational algorithm

The algorithm used here is similar to the algorithm outlined in Section 2.1, but with an important difference. In the previous algorithm, individuals could simply be counted up and the species treated as a whole, with the competitiveness of the species simply proportional to the number of individuals. Now, however, individuals must be treated individually, as they all reside in different locations, which gives each of them different competitiveness.

The following steps describe the new algorithm:

1. At the beginning of the simulation, a set of calculations are performed that evaluate the distance between any two sites. This is then stored for future use.

2. When a species is created, the fecundity function for that particular species is evaluated for all sites in the lattice and stored for future use.

3. Additionally, when a species is created, the dispersal kernel for that particular species is evaluated for all possible distances between sites in the lattice and stored for future use

4. At the beginning of each time step, a copy is made of the state of individuals present in the lattice. This copy is then held static until the beginning of the next time step.

5. The number of individuals of each species is added up and stored.

6. Every site on the lattice is then visited. The following steps happen at each site:

   (a) The individual occupying the site from the previous generation is removed.

   (b) Each individual 'competes' to be the one whose offspring establishes at the newly vacant site. A weighted probability of establishment is calculated for each individual. This is given by the following equation:

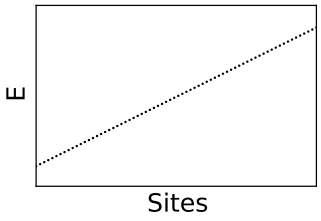
$$\text{Weighted prob} = F(\mu, \sigma; E) \times D_{\text{norm}} \tag{4.6}$$

   (c) All weighted probabilities from all individuals are added up and normalised such that they sum to 1.

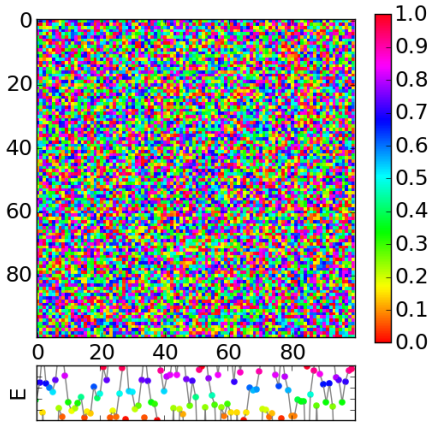
   (d) The following algorithm is then applied:

PROBS = array containing the weighted probability of each individual
INDICES = array containing the identifying index of each individual
A = random number between 0 and 1 picked from a uniform distribution
i = 0
COUNT = PROBS[0]
loop
  if COUNT $\geq$ A **then**
    return INDICES[i]
  else
    COUNT = COUNT + PROBS[i]
    i+=1
  end if
end loop

   (e) The identifying index of the winning individual has been selected, and a copy of that individual is placed in the site.



(a) *Illustration of resource spectrum*



(b) *Arrangement of sites in space*

As a new probability must be calculated for every individual at every time step, and on average $\frac{N}{2}$ calculations are needed in order to select an individual, the computational complexity of this algorithm scales nonlinearly as N^3.

4.2 An uncorrelated environment using a uniform heterogeneous resource spectrum

As the model is now spatial, the arrangement of sites becomes important. A spatial environment is made up of an underlying resource spectrum, which defines all sites found in the environment. These sites are then arranged on a 2D lattice so that the environment has spatial structure.

All environments in this chapter use the uniform heterogeneous resource spectrum introduced in Section 2.2 unless stated otherwise. This resource spectrum has N unique sites that are spaced at equal distance in E space from one another, as illustrated in Figure 4.3a

These sites are then placed randomly on a 2D lattice. Figure 4.3b shows an example of such an arrangement, where sites are colour-coded by their E value. Underneath the plot, a 1D slice of the environment is shown for y=50, showing the spatial location of sites against their E value. We initially use uncorrelated environments such as this one to explore the effects of limiting dispersal.

4.3 Fixed δ, variable σ in an uncorrelated environment

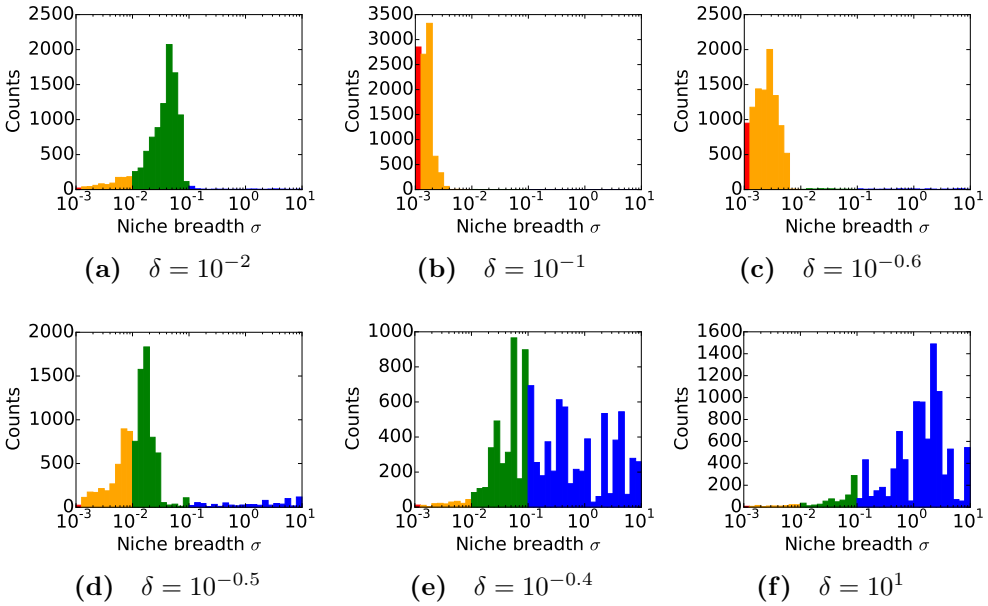
To explore the effects of limited dispersal on the specialist-generalist balance, we now run a number of simulations in which δ is held constant and σ is varying. This addresses the question of whether a species should take the strategy of being a specialist or a generalist in environments in which long-range dispersal is easy such as the open ocean [38], or environments in which long-range dispersal is more difficult. Simulations are run for 5000 time steps and one replicate is taken of each parameter set. We chose 5000 rather than the 30000 time steps used previously because simulations that include spatial structure are far more computationally intensive than well-mixed simulations and to run a large number of simulations for a longer time would have been impractical as each simulation for 5000 time steps took, on average, two days to run. Nonetheless, we are confident that running for this shorter time results in a qualitatively accurate picture of the community composition at steady state. As before, we analyse the end state of the simulation, taking data from the final 500 time steps.

Results

Figure 4.4 shows simulation results for the geometric mean value of the niche breadth σ for a range of fixed simulations, each with fixed dispersal parameter δ. The full distribution of specialist-generalist strategies (σ value) is plotted for some of these simulations in Figure 4.3.

Our results show that for a system that has high dispersal (δ between 10^0 and 10^2), the system tends to a generalist-dominated state. This state does not appear to change in either its geometric mean σ or Shannon diversity with changes in δ over this range of δ values. Figure 4.3f shows a typical σ distribution for this case, which is qualitatively the same as seen earlier in Chapter 2 for simulations with infinite dispersal resulting in the generalist-dominated state (Figure 2.8b).

Values of δ below 10^0 favour a larger population of mild specialists (Figure 4.3e), while the generalist population is maintained. However, this state only exists for a narrow range of δ values. Between $\delta = 10^{-0.4}$ and $\delta = 10^{-0.6}$, the system undergoes a transition (Figures 4.3e, 4.3d and 4.3c) in which it becomes more and more specialist-dominated with decreasing δ, with a corresponding increase

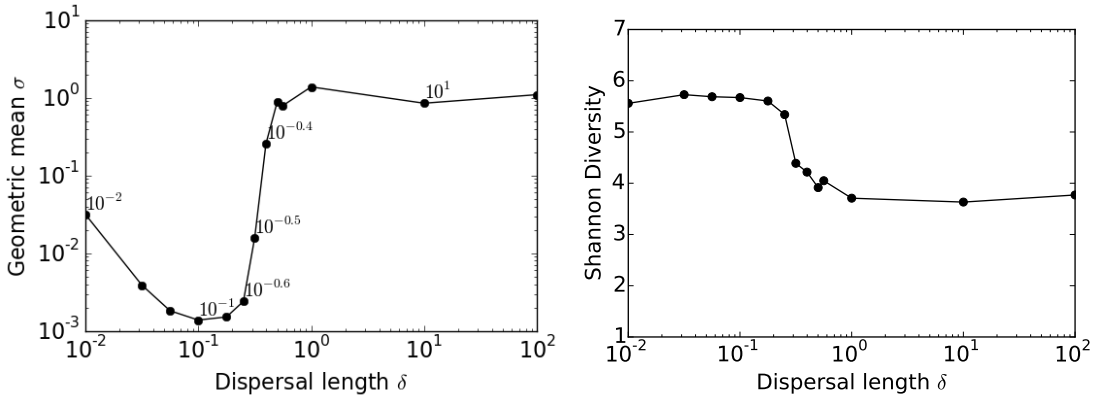


(a) $\delta = 10^{-2}$ **(b)** $\delta = 10^{-1}$ **(c)** $\delta = 10^{-0.6}$

(d) $\delta = 10^{-0.5}$ **(e)** $\delta = 10^{-0.4}$ **(f)** $\delta = 10^{1}$

Figure 4.3 *σ distributions occurring for some values of δ.*

Colour coding shows σ categories: blue: generalist; green: mild specialist; yellow: intermediate specialist; red: extreme specialist. We can see that there is a dramatic shift in community composition between $\delta = 10^{-0.4}$ and $\delta = 10^{-0.6}$.



(a) *Geometric mean σ* **(b)** *Shannon Diversity*

Figure 4.4 *Geometric mean σ and Shannon diversity for a range of fixed δ in an uncorrelated environment. Labelled points indicate that a corresponding σ distribution is shown in Figure 4.3.*

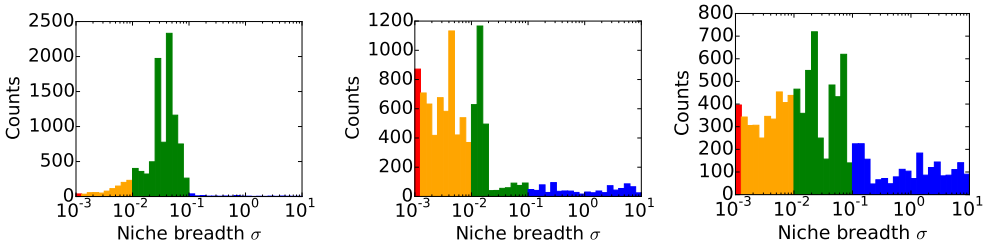
in Shannon diversity. At $\delta = 10^{-1}$ the system is dominated by extreme specialists (Figure 4.3b).

However, as δ is decreased further, we observe that the system actually becomes less specialist-dominated. For the extremely low dispersal value of $\delta = 10^{-2}$, we see from Figure 4.3a that the system predominantly consists of mild specialists.

We can draw a number of interesting observations from these results. Firstly, we note that generalists apparently cannot survive in our model when dispersal is limited, i.e., for $\delta < 10^{0}$. This is most likely because the generalist strategy relies on having access to a large number of sites with a feasible possibility of winning some of them, which is not possible if dispersal is limited. Even when dispersal is very limited (i.e., for $\delta = 10^{-2}$), there are still no generalists that can survive.

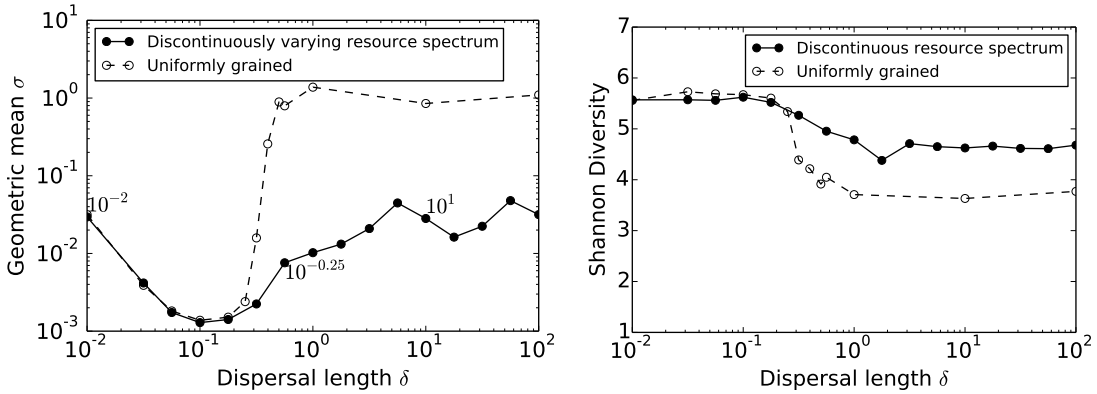
We speculate that for specialists, being able to link amenable sites together in 'corridors' is more important than accessing them all at once. A specialist occupying an attractive site in a medium-range dispersal community is very unlikely to displaced from its site by a generalist, as it only has serious competition from individuals located close to it and is safe from competition pressure from generalists located further away. As it exists in relative safety from being displaced from its site, there is no rush to establish a large population quickly. Instead, its population can grow slowly, as long as attractive sites exist within accessibly reach of one another. In a medium-range dispersal community (around $\delta = 10^{-1}$), a species has a rough dispersal radius of around 23 sites, which gives it access to around 16% of the sites in the habitat (Figure 4.2). An extreme specialist, however picky, has a reasonable chance of finding an attractive site in this area. As it is highly specialised to that site, it has a good chance of displacing whatever less specialised individual might be occupying it. Once the species becomes established on the second site, it then has access to the 16% of sites surrounding *that* site. Using this strategy, an extreme specialist species may spread through the habitat slowly. Hence, we see a highly specialised population emerging when $\delta = 10^{-1}$ (Figure 4.3b).

However, if dispersal ranges are too small, an extreme specialist is unlikely to be able to link amenable sites together. If $\delta = 10^{-2}$, the dispersal radio of each species is around 3.7, giving it access to only around 0.5% of the sites in the habitat. However, as it is less picky, a mild specialist is far more likely to find amenable sites in this small sample. We therefore find a population made up predominantly of mild specialists at such short dispersal ranges (Figure 4.3a).



(a) $\delta = 10^{-2}$ **(b)** $\delta = 10^{-0.25}$ **(c)** $\delta = 10^{1}$

Figure 4.5 *σ distributions occuring for different δ values in environments with discontinuously varying resource spectrum*



(a) *Geometric mean σ* **(b)** *Shannon diversity*

Figure 4.6 *Geometric mean σ for a range of fixed δ in an environment with a discontinuous resource spectrum (ρ=0.2), with an environment with uniformly grained resource spectrum shown for comparison. σ distributions of labeled simulations shown in Figure 4.5*

However, as this strategy is still fundamentally a specialist strategy, we do not see any generalists succeeding in this short-range environment.

4.3.1 Results: Using a discontinuously varying resource spectrum with ρ=0.2

In Chapter 2, we established a model with a discontinuously varying resource spectrum, in which the severity of discontinuities in the resource spectrum varied with a parameter ρ (Section 2.4.2). We found that for intermediate values of ρ, there was wide coexistence between generalists and specialists (Figure 2.14b). We now wish to explore whether this wide coexistence of generalists and specialists can be found when dispersal is limited in an uncorrelated environment with this resource spectrum.

Using uncorrelated environments with an underlying discontinuously varying resource spectrum with $\rho = 0.2$, we once again run simulations for a range of fixed δ for 5000 time steps to find the resultant specialist-generalist balance.

Simulation results for the geometric mean σ for simulations at fixed δ with the corresponding Shannon diversity are plotted in Figure 4.6, with the geometric mean σ and Shannon diversity from environments using a uniformly grained resource spectrum from the previous section plotted in white for comparison. Individual σ distributions for some of these simulations are plotted in Figure 4.5.

We see that for high values of δ, the two systems behave differently. For values of δ that correspond to a generalist-dominated regime in the previous section, in this case coexistence is seen between generalists and specialists (Figure 4.5c), with σ distribution similar to that seen in the well-mixed model with a discontinuously varying resource spectrum (Figure 2.14b). This suggests that when dispersal lengths are sufficiently long ($\delta > 10^0$), the resulting σ distribution is close to that seen for the well-mixed case for that resource spectrum.

However, for lower values of δ, we see similar results to those seen in the previous section, with the two different types of resource spectra giving near-identical results for $\delta < 10^{-1}$. This shows that when dispersal is limited, the underlying resource spectrum of the whole environment has much less of an effect.

4.4 Fixed σ, variable δ in an uncorrelated environment

We now turn to the question of whether species which already have a specialist or generalist strategy should invest in short or long range dispersal abilities. Therefore, we performed simulations with different values of fixed σ, while δ varies. Once again, we run simulations on an uncorrelated environment with a uniform resource spectrum, and simulations are run for 5000 time steps, using data from the final 500 time steps.

Similar to our separation of σ into different categories, here we categorise the dispersal ranges of our species into three categories: long (LD, $10^{-0.5} < \delta < 10^2$), medium (MD, $10^{-1.5} < \delta < 10^{-0.5}$) and short (SD, $10^{-2} < \delta < 10^{-1.5}$) range dispersers (Figure 4.7). These categories and colour coding will be referred to

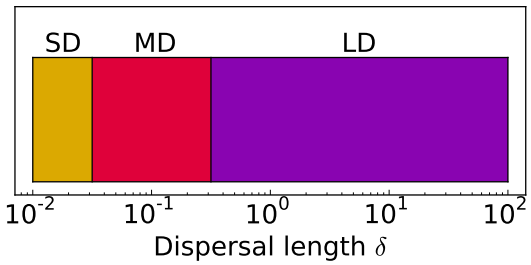


Figure 4.7 *Categories of δ*

As with σ categories, species are separated into categories according to their δ value. Once again, the boundaries between categories are somewhat arbitrary.

throughout the rest of this chapter.

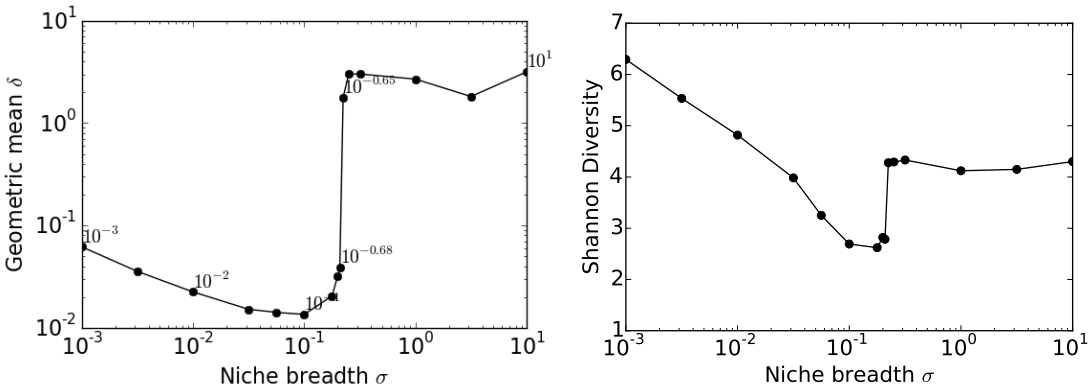
Results

Figure 4.8 shows the geometric mean δ and Shannon diversity for a range of simulations performed with fixed σ. As we did with the σ distributions, we can examine the δ distributions at various points (Figure 4.9).

For high values of σ ($10^{-0.65} < \sigma < 10^{1}$, corresponding to almost the entire generalist range), the resulting geometric mean δ is high, consisting of a community of long-range dispersers and some medium-range dispersers (Figures 4.9f and 4.9e). The Shannon diversity does not change in this regime, and is close in value to the Shannon diversity seen in the high-δ resulting in high σ regime of the previous section (Figure 4.4b).

Between $\sigma = 10^{-0.65}$ and $\sigma = 10^{-0.68}$ a sharp transition takes place. For values of σ larger than this threshold, long and medium range dispersers coexist; for smaller values of σ only a very narrow 'band' of short and medium range dispersers can exist (Figure 4.9d). As σ decreases, the community becomes more short range, reaching a minimum of geometric mean δ at $\sigma = 10^{-1}$ (Figure 4.9c). This transition is accompanied by a sharp drop in Shannon diversity. In this regime (approximately $10^{-1.25} < \sigma < 10^{-0.68}$) the community is made up of a small number of high-abundance species on the border between mild specialism and generalism, which disperse only very small distances.

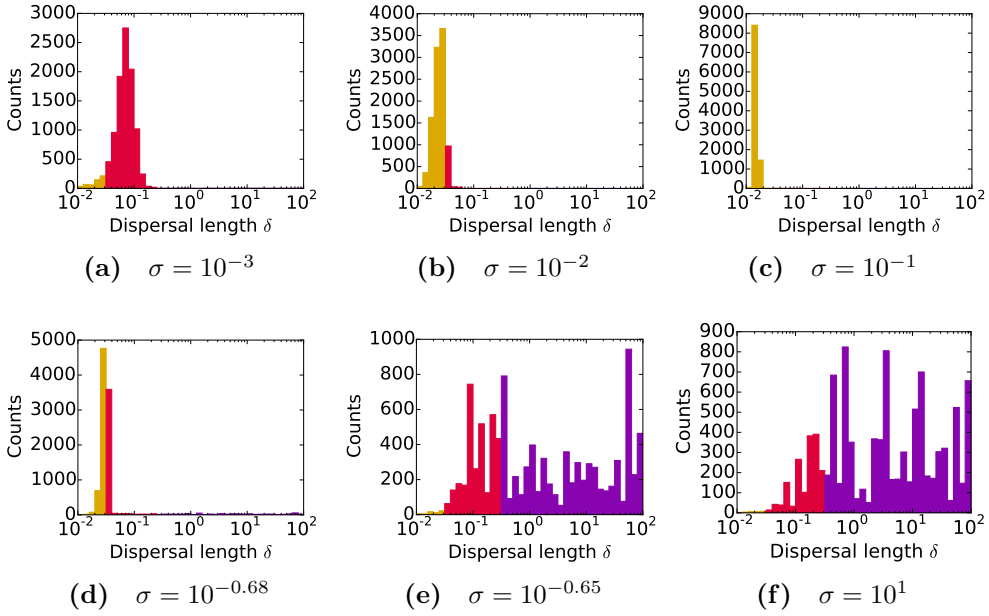
As σ decreases further, the geometric mean δ increases again (Figures 4.9b and 4.9a), although the community is still composed of short and medium range



(a) *Geometric mean δ* **(b)** *Shannon Diversity*

Figure 4.8 *Geometric mean δ and Shannon diversity for a range of fixed σ*

We see that a community populated by generalists tends to result in species with long-range dispersal abilities. A transition occurs around $\sigma = 10^{-0.68}$*, where the dispersal ability of the resultant community becomes much lower. This is accompanied by a drop in Shannon diversity. As σ is decreased further, the dispersal ability rises again, with a corresponding large rise in Shannon diversity.*



(a) $\sigma = 10^{-3}$ **(b)** $\sigma = 10^{-2}$ **(c)** $\sigma = 10^{-1}$

(d) $\sigma = 10^{-0.68}$ **(e)** $\sigma = 10^{-0.65}$ **(f)** $\sigma = 10^{1}$

Figure 4.9 *Resultant δ distributions for communities composed uniformly of species with different σ values.*

We can observe a large difference in resultant δ distribution between $\sigma = 10^{-0.65}$ *and* $\sigma = 10^{-0.68}$*.*

dispersers, with no long-range dispersers. This is accompanied by a steady increase in Shannon diversity with decreasing σ, as we expect niche differentiation to increase with a decrease in niche breadth. We note that even when short-range dispersers are most successful, at $\sigma = 10^{-1}$, extremely short range dispersers ($\delta < 10^{-1.9}$) do not appear. Such a species, for example with $\delta = 10^{-2}$, has an effective dispersal distance of around 3.5 lattice sites. It would therefore appear that in our model it is important for the survival of a species to be able to disperse at least some distance.

Comparing these results to the results from the previous section in which δ was fixed and σ varied (Figures 4.4 and 4.3), we see many parallels. Comparing plots for geometric mean (Figures 4.4a and 4.8a), we see that a community with a high σ will result in a community with a high δ, and vice versa, suggesting that long-range dispersal is closely associated with generalism in our model. Likewise, intermediate σ values are associated with short dispersal lengths, and low σ values are associated with medium dispersal lengths.

4.5 Correlated environments

Up until now, we have been using uncorrelated, randomly shuffled environments. However, resources in nature are often correlated in space, for example, arable fields of different crops. Therefore, we now wish to produce correlated environments in our model.

In order to arrange the environmental sites in a correlated way, we use a swapping algorithm. First, an array of sites with a uniform heterogeneous resource spectrum is randomly shuffled. The swapping algorithm described below is then applied. The aim of this algorithm is to decrease the 'disorder' of the system, where disorder is high if there are many sites of very different E values close together. For example, the disorder of the uncorrelated environment shown in Figure 4.3b is very high.

The local disorder, S_{xy} of a site located at (x,y) in the lattice is defined by summing up the difference between the E value at this site and the E values of sites within a distance d of the site. This difference is then raised to a power m

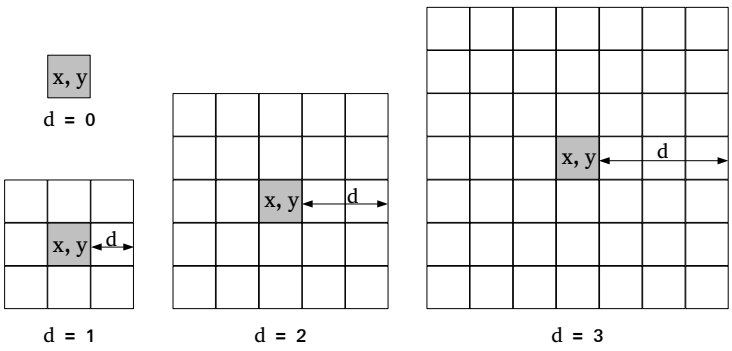


Figure 4.10 *The lattice sites that contribute to the disorder of site i*

The larger d is, the larger the number of surrounding sites that contribute to the effect on the central site.

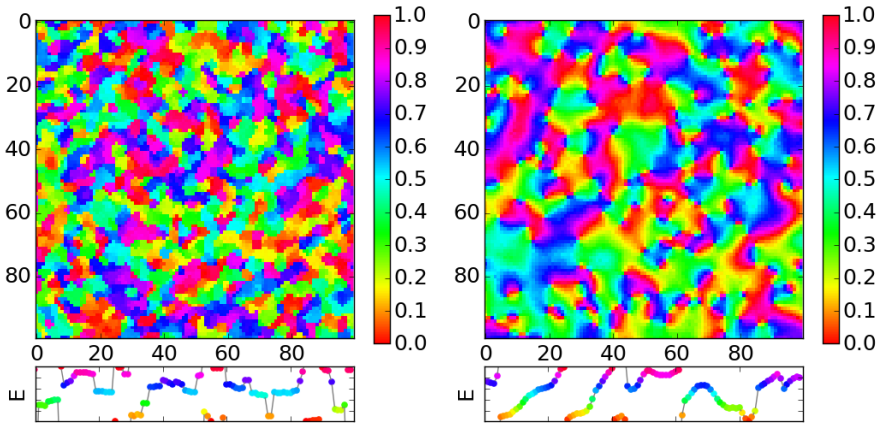
(Equation 4.7).

$$S_{x,y} = \sum_{i=x-d}^{x+d} \sum_{j=y-d}^{y+d} \Delta(E(i,j) - E(x,y))^m \tag{4.7}$$

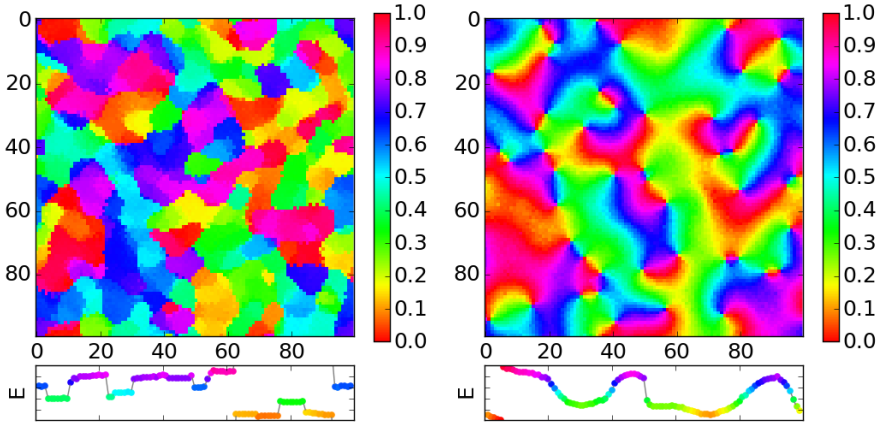
To arrange the environment, we randomly pick pairs of sites and calculate the local disorder at each site. We then calculate the hypothetical local disorder that these sites would have if they were swapped. If the total hypothetical local disorder is lower than the total local current disorder, the positions of these sites are swapped. This process is then repeated 3 million times.

Figure 4.11 shows a range of environments that can be made by setting d equal to 1, 2 or 3, and setting m equal to either 0.5 or 2. We can see that increasing d results in the domains in the environment becoming larger, while m controls the smoothness of the environment. For example, setting m=0.5 (Figures 4.11a, 4.11c and 4.11e) results in an environment in which domain borders are sharp and discontinuous, while setting m=2 (Figures 4.11b, 4.11d and 4.11f) results in an environment in which domains merge smoothly into one another without any sharp boundaries. These correlations are short-ranged, having a correlation length of around 0.05 of the total length scale of the habitat in the case of d=1 and of around 0.16 in the case of d=3.

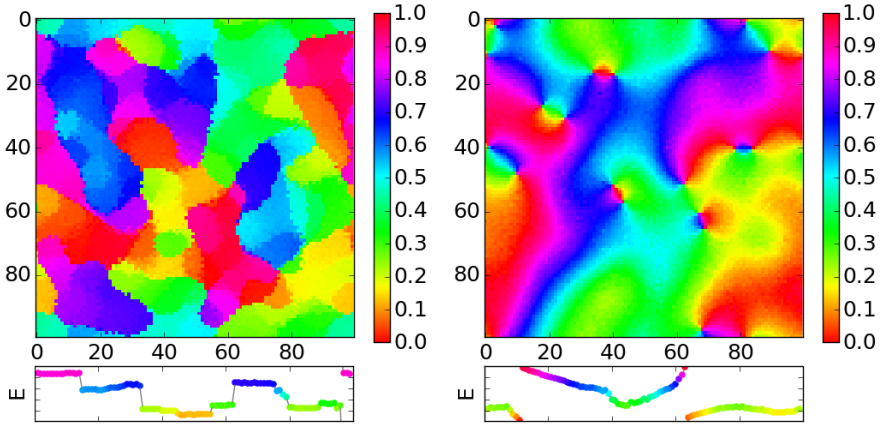
These two different types of environment can be thought of as being ecotonal (m=0.5), with sharp discontinuities between different types of habitat, and ecoclinal (m=2), in which differences between habitat vary smoothly in



(a) $d = 1,\ m = 0.5$ **(b)** $d = 1,\ m = 2$



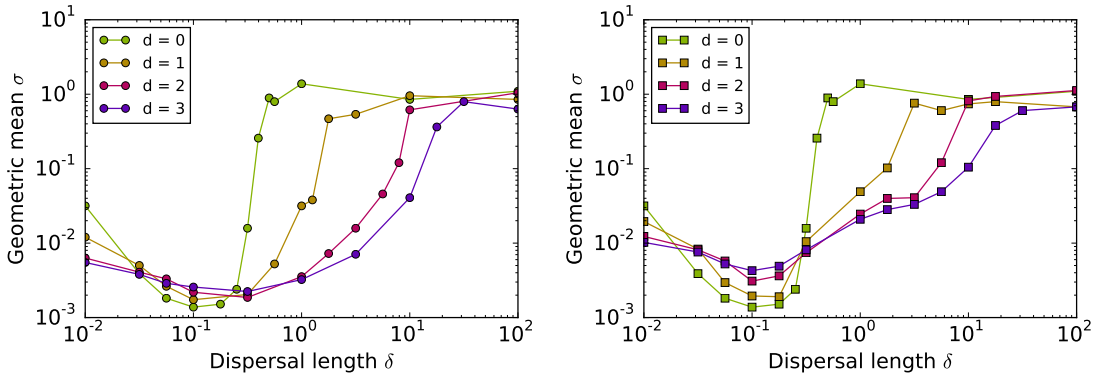
(c) $d = 2,\ m = 0.5$ **(d)** $d = 2,\ m = 2$



(e) $d = 3,\ m = 0.5$ **(f)** $d = 3,\ m = 2$

Figure 4.11 *Examples of environments formed by a range of d and m.*

The lower plots show the changing of the environment along the x axis at y=50. d controls the size of the domain areas, while m controls how smoothly transitions between domains occur.



(a) $m = 0.5$ **(b)** $m = 2$

Figure 4.12 *Geometric mean σ for a range of fixed δ in environments with a range of d with m=0.5 (a) and m=2 (b)*

space. It is important to note that although the *spatial organisation* of these environments is different, the underlying *resource spectrum* is the same; the uniform heterogeneous resource spectrum introduced in Figure 2.2b. We also note that an environment with d=0 refers to an unsorted (randomly shuffled and so uncorrelated) environment (Figure 4.3b).

4.6 Fixed δ, varying σ on correlated environments

We now investigate the effects of the arrangement of resources in space on the specialist-generalist balance in communities with fixed δ. Once again, we run simulations for a range of fixed δ for 5000 time steps, for each combination of d=1,2 and 3, and m=0.5 and 2.

Figure 4.12 shows the geometric mean σ for all types of environments. Note that the environments used are not the *specific* environments illustrated in Figure 4.11, but are environments with similar properties that have arisen from the same input parameters; indeed, each environment is different for each simulation. The environments are split into m=0.5 (ecotonal environments, Figure 4.12a) and m=2 (ecoclinal environments, Figure 4.12b). The coloured lines show results for environments with different values of d.

Examining the results for larger values of d, we see a similar pattern as we saw for the uncorrelated case where d=0, for both $m = 0.5$ and $m = 2$. As d increases, the transition to a more specialist dominated state happens at a higher value of

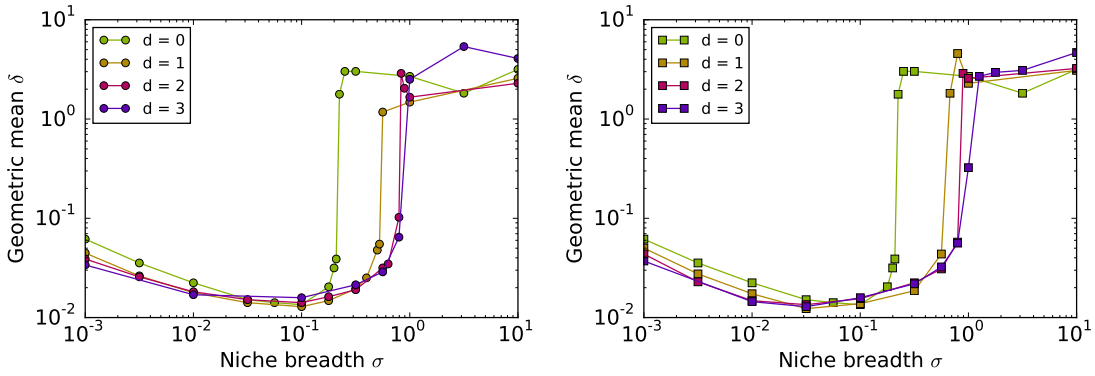
δ. However, interestingly we note that the lower d is, the lower the minimum value of geometric mean σ is. This is true for both m=0.5 and m=2, but is more pronounced in the case of m=2. This may be because the possible extreme specialist strategy discussed in Section 4.3, in which extreme specialists slowly disperse through the environment between isolated attractive sites, does not work as well when environments are correlated. In an uncorrelated environment, extreme specialists do not face much competition from specialists. Ideal sites are inconveniently spread out for all specialists, and in our simulations all species disperse the same distance. In this unideal situation, extreme specialists can prosper slowly by being the most successful on each site. However, if the environment is correlated, intermediate specialists may be at more of an advantage, as many sites within their reach may be amenable. An intermediate specialist species may then grow its population in that area, crowding out any extreme specialists. This effect is seen more in ecoclinal (m=2) than ecotonal (m=0.5) because in ecotonal environments, more similar sites are gathered closer in space than in ecoclinal environments (compare Figures 4.11e and 4.11f). This may allow more specialised species to establish themselves in a local area when similar sites are clumped, making the effect less pronounced.

Examining the differences between the ecotonal environment (Figure 4.12a) and the ecoclinal environment (Figure 4.12a), we see that broadly, the ecotonal environment results in communities being more specialist than the ecoclinal environment, although this difference is only seen for d=2 or 3. In both of these cases, the minimum value of σ is lower in the ecotonal environments, and the transition from generalism to specialism is steeper, resulting in more specialist communities in the range $10^{-0.5} < \sigma < 10^{1}$.

4.7 Fixed σ, varying δ on correlated environments

We now investigate the effects of the environment on pre-defined specialists and generalists forming a dispersal strategy. In these simulations, σ is held constant on a number of correlated environments and σ varies.

Examining Figure 4.13, we see that more correlated environments result in the transition from long range to short range dispersers occurring at a higher value of σ (more generalist). This can be explained by more generalist species having an incentive to shrink their dispersal range when large amounts of their resources



(a) $m = 0.5$ **(b)** $m = 2$

Figure 4.13 *Geometric mean δ for a range of fixed σ in environments with a range of d with m=0.5 (a) and m=2 (b)*

are in one place, as they have no need to disperse further.

We do not see any significant difference between ecoclinal and ecotonal environments in this case.

4.8 Conclusions

Our models show that the trade-off between specialism and generalism, and the trade-off between long and short range dispersal are closely linked.

As previously discussed in Section 1.4.3, it is unclear whether specialists should favour short range dispersal or long range dispersal. As specialists need specific types of habitat, they may have to disperse a long distance in order to find those habitats, which may result in long-range dispersal. However, if they are in an ideal habitat already which is large enough to accommodate them, there may be no reason to attempt to find more of it, which may result in short-range dispersal.

Our results show evidence of this trade-off in the model. When niche breadth σ is held constant and δ varies, we see that specialists ($10^{-3} < \sigma < 10^{-1}$) use short- or medium-range dispersal strategies in all types of environment (Figure 4.13). However, mild specialists use a shorter dispersal range than intermediate or extreme specialists. This suggests that in this model, specialists use as short a dispersal distance as they can while still having access to favourable sites. Specialists with a wider niche are less constrained in their habitat choice and will find a wider range of habitat favourable, meaning they do not need to disperse

as far to find it. Specialists with a narrow niche are more constrained, and so may have to travel further to find their ideal habitat. In addition, environments with longer-range correlation (higher d) result in a lower average dispersal range in extreme and intermediate specialists, as in these environments, a large amount of similar habitat is spatially close together, so a specialists may disperse less distance to find it.

In addition, our results suggest that different specialist strategies may work depending on the correlation of the environment. If the environment is uncorrelated, highly specialised specialists that can disperse far enough to link non-adjacent attractive sites may be successful. If the environment is correlated, less specialised specialists that are able to colonise a large local area may be successful. This is suggested by the paradoxical result that, in a community with intermediate dispersal ranges (Figure 4.12), extreme specialists are successful in an uncorrelated environment, while intermediate specialists are successful in a highly correlated environment.

However, it is important to note that the environments used here do not change in time. A temporally heterogeneous environment in which could encourage more long-range dispersal in specialists, as they would be forced to move from their sites to find their ideal resources.

Chapter 5

Analysis of Sequence Data

5.1 Introduction

This chapter details a collaboration with the Andrew Free lab in the School of Biological Sciences. In the previous chapters, we took a theoretical and idealised view of community ecology, and the factors that shape a community of species living together in the same location. In this chapter, we instead look at some real data, and examine the factors that cause microbial species to form a community. The previous chapters showed that entire communities may be co-selected and a small stochastic event in the right place at the right time may lead to a community-wide shift in composition. Here, we show that the same can be true of real microbial communities.

In the previous chapters we saw that a diverse community of specialists can coexist by the mechanism of niche partitioning, in which each specialist utilises a different type of environmental resource to avoid competition with one another. We also saw that a community of generalists can coexist by being able to utilise a range of different resources. However, for specialists in a homogeneous environment, both theory and experiment show a single clear winner will emerge and diversity will collapse. Generalists in a homogeneous environment, however, will continue to have a diverse community due to neutral dynamics.

We now explore these ideas in a real ecological system in two studies, using micro-organisms sampled from Blackford Pond, a local pond in Edinburgh. In the first study, microorganisms are exposed to a range of added cellulose and

sulphate concentrations and left to develop in a microcosm environment for several months. Analysing the communities, we can establish which microbes are generalists (which are unaffected by the different conditions), and which function as specialists (which are highly affected by the changing conditions).

In the second study, a number of identical communities are set up in microcosms and left to develop, again for several months. We find that different communities of microbes emerge in different microcosms. Correlations can then be found between microbes that co-occur across different communities, suggesting functional interactions. We can then establish whether generalists or specialists tend to interact with one another.

In the next two subsections, we outline some biological terminology and experimental techniques that will be useful for understanding the rest of the chapter.

5.1.1 Taxonomy in microbial ecology

Every organism has its own unique genetic sequence within its DNA, made up of combinations of the four nucleobases: Adenine (A), Guanine (G), Cytosine (C) and Thymine (T). Members of the same species have a very similar, but not identical, genetic sequence. As organisms become more and more unrelated, their DNA sequence differs more and more. Assigning taxonomy is an attempt to divide the level of relatedness between organisms into different levels.

There are eight main levels of taxonomy; Domains, Kingdoms, Phyla, Classes, Orders, Families, Genera and Species. As an example, we will lay out the taxonomy of the domestic cat:

- Species *catus*
- Genus *Felis* (includes other small cat species such as desert cats)
- Family *Felidae* (includes all big and small cats such as tigers and lions)
- Order *Carnivora* (includes many species of placental mammals including bears, dogs and weasels)
- Class *Mammalia* (includes all mammals)
- Phylum *Chordata* (includes all vertebrates)

- Kingdom *Animalia* (includes all animals)
- Domain *Eukaryota* (comprises all cells with a nucleus and includes plants and fungi as well as animals)

The taxonomy of an individual refers to the names of all the taxonomic levels it belongs to, although in common parlance species are generally referred to only by their genus and species, for example *Felis catus* or *Homo sapiens*. A taxon is a group of organisms that are similar enough to each other to be considered as a unit. In most macroscopic cases, 'taxon' is interchangeable with 'species', but in the microbial world this is not always true, for reasons that will be explained later.

All microorganisms of concern to this work belong to the domains Archaea and Bacteria, which together form the prokaryotes, which are cells without a nucleus. Just like the cat, these microorganisms have levels of taxonomy that can be used to determine their function, and similarity to one another.

5.1.2 Identifying microbial taxa using 16S rRNA gene sequencing

Until the 1980s, the only way of determining what organisms were present in a microbial sample was to culture them (allow them to grow on an agar plate), and then examine the resultant culture under a microscope, or examine which nutrients allow them to grow [120]. This is very time-consuming and only allows a small amount of a complex microbial community to be detected. Modern methods for determining the organisms in a microbial sample now depend on DNA sequencing.

'DNA sequencing' is the act of finding the sequence of the DNA code (the pattern of A, G, C and T nucleobases) that codes for an organism. These nucleobases come in pairs, which bond to one another using hydrogen bonds. A always binds to T, and G always binds to C. A long string of these pairs, curving into a helical shape, forms the famous double helix. Each strand of the double helix contains the information needed to recreate the full helix, as a C must always be matched with a G, and so on. A 'gene' refers to a part of the genetic code that encodes a particular protein molecule or RNA.

Sequencing the entire genome of an organism is costly and time consuming.

Therefore, if the objective is to determine the abundances of many organisms in a microbial community sample, it is far more effective to identify a single gene present in all of the organisms of interest and sequence only that, using it as a marker. Because different organisms differ slightly in the exact sequence of this gene, the sequence fragments can be used as markers for different organisms. This can then be used to compare the genes found against taxonomic information already known and is known as 'metataxonomics' [100].

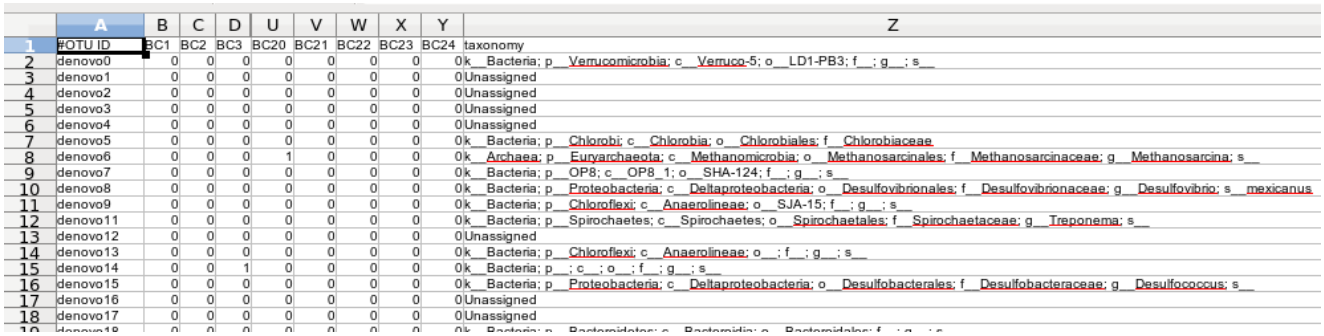
The most widely used gene for this purpose is the 16S ribosomal RNA (16S rRNA) gene. This gene is present in all prokaryotes and codes for part of the ribosome, an essential part of the cell machinery involved in protein production [120]. Each copy of the gene in different taxa contains small differences, such that a taxon can be identified by the exact form of its 16S rRNA gene.

In order to sequence these genes, the abundance of the DNA fragments must be massively amplified. To do this, a technique called PCR (Polymerase Chain Reaction) is performed on the samples[1]. These samples are then put through a sequencing machine to return a large (in the order of 1,000) set of sequences associated with each sample. It can generally be assumed that the number of sequences of a particular gene sequence variant will correspond roughly to the number of organisms in the original sample that held that gene, although due to some issues with the PCR technique and also because some organisms have multiple copies of the 16S gene in their genomes, this will not be an exact match.

Due to the fast rate of evolution in microorganisms, the divide between different taxonomic levels is not particularly clear cut. For this reason, it is useful to talk about 'Operational Taxonomic Units' (OTUs) rather than 'species'. An OTU is simply a group of sequences of high enough similarity to be considered a unit. QIIME (Quantitative Insights Into Microbial Ecology) [31] is a bioinformatics software used to cluster sequences into groups that are above a specified similarity to one another, and classify the group as a single OTU.

All sequences within QIIME are compared for similarity in their genetic code

[1]PCR is made up of three stages. First, double stranded DNA is separated into single stranded DNA. Second, oligonucleotide primers are added and anneal to complementary binding sites in on the single-stranded DNA. The primer is specific to the gene that is to be amplified, in this case the 16S rRNA gene. Thirdly, nucleotides (building blocks of DNA) are added and extend from the primer, forming the complete complementary strand of the gene in question. This newly formed double-stranded fragment of DNA is then separated into single stranded DNA and the process starts again. By going through many cycles, fragments of the gene are replicated many times, resulting in millions of copies. The huge amplification of the DNA allows for sequencing [120].



	A	B	C	D	U	V	W	X	Y	Z
1	#OTU ID	BC1	BC2	BC3	BC20	BC21	BC22	BC23	BC24	taxonomy
2	denovo0	0	0	0	0	0	0	0	0	k__Bacteria; p__Verrucomicrobia; c__Verruco-5; o__LD1-PB3; f__; g__; s__
3	denovo1	0	0	0	0	0	0	0	0	Unassigned
4	denovo2	0	0	0	0	0	0	0	0	Unassigned
5	denovo3	0	0	0	0	0	0	0	0	Unassigned
6	denovo4	0	0	0	0	0	0	0	0	Unassigned
7	denovo5	0	0	0	0	0	0	0	0	k__Bacteria; p__Chlorobi; c__Chlorobia; o__Chlorobiales; f__Chlorobiaceae
8	denovo6	0	0	0	1	0	0	0	0	k__Archaea; p__Euryarchaeota; c__Methanomicrobia; o__Methanosarcinales; f__Methanosarcinaceae; g__Methanosarcina; s__
9	denovo7	0	0	0	0	0	0	0	0	k__Bacteria; p__OP8; c__OP8_1; o__SHA-124; f__; g__; s__
10	denovo8	0	0	0	0	0	0	0	0	k__Bacteria; p__Proteobacteria; c__Deltaproteobacteria; o__Desulfovibrionales; f__Desulfovibrionaceae; g__Desulfovibrio; s__mexicanus
11	denovo9	0	0	0	0	0	0	0	0	k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__SJA-15; f__; g__; s__
12	denovo11	0	0	0	0	0	0	0	0	k__Bacteria; p__Spirochaetes; c__Spirochaetes; o__Spirochaetales; f__Spirochaetaceae; g__Treponema; s__
13	denovo12	0	0	0	0	0	0	0	0	Unassigned
14	denovo13	0	0	0	0	0	0	0	0	k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__; f__; g__; s__
15	denovo14	0	0	1	0	0	0	0	0	k__Bacteria; p__; c__; o__; f__; g__; s__
16	denovo15	0	0	0	0	0	0	0	0	k__Bacteria; p__Proteobacteria; c__Deltaproteobacteria; o__Desulfobacterales; f__Desulfobacteraceae; g__Desulfococcus; s__
17	denovo16	0	0	0	0	0	0	0	0	Unassigned
18	denovo17	0	0	0	0	0	0	0	0	Unassigned

Figure 5.1 *Example of a rarified OTU table*

A screenshot of a small sample of a rarified OTU table. BC1–BC24 are samples and denovo0–denovo17 are taxa with taxonomic identify assigned if available. The identity of many taxa is unknown and their taxonomy is thus 'Unassigned'. Of those that are known, most do not have full taxonomic information. As this table is rarified, much of the original abundance information has been removed, resulting in many taxa with zero relative abundance across all samples.

and if a group of sequences are above the specified threshold of similarity (the standard to use is 97%), they are treated as a single 'taxon' for the purposes of all henceforth analysis. The most abundant sequence in an OTU is then selected and its taxonomy is assigned, if available, from a database of known sequences. All sequences in the OTU are then assumed to have the same taxonomy. In practice, complete taxonomic information all the way down to species is typically known for only very few sequences within a given sample. The most detailed level of taxonomy for most sequences is only at genus, family or order level, and a large number typically have no taxonomic information at all, as there are still many microbial species that are unknown. An OTU abundance table is then created which shows the abundance (number of sequences) of each OTU in each sample, with whatever taxonomic information is available for that OTU. Figure 5.1 shows a simplified example of an OTU table.

The total abundances of OTUs that result from sequencing a given sample may vary widely from sample to sample, and give no reliable indication of the genuine abundances in each sample. This is due to random variation in the sequencing process. To ensure that all samples have an equal influence on the results, the data must either be normalised, so each abundance becomes the fractional abundance of all OTUs seen in that sample, or rarified, in which data is randomly deleted such that all samples have the same number of remaining sequences. In the first experiment we describe here the datasets have been rarified, and in the second

they have been normalised.

5.2 Functional group analysis of an experiment varying cellulose and sulphate in Winogradsky columns

The aim of this experiment was to determine the effect of perturbing nutrient concentrations on the composition of a microbial community. We are interested both in the microbial community as a whole, and different 'functional groups' within the community. A microbial functional group is a subset of the community that performs a specific role in that community - for example, the degrader group performs the important task of degrading complex organic molecules into smaller molecules so that they can be used as a nutrient source by other taxa.

A rough picture of a microbial community can be obtained by examining its total abundance (number of sequences detected) and diversity (roughly, the range of sequences detected). Although we cannot obtain a measure of the true abundance of a community by this technique, we can find the relative abundance of functional groups within that community and examine the diversity within those groups.

An interesting question to ask about a functional group is whether or not it is stable to environmental perturbations.

A stable group will continue to perform its function in the community through a range of environmental conditions, even if the composition of individual taxa within the group changes. We expect to see a relatively unchanged value of abundance and diversity for that group across diverse environmental conditions.

A stable functional group could arise because the group is composed of generalist species, which are able to continue to function in different environmental conditions by using alternative metabolic pathways (in other words, they are functionally redundant [99][56]), or because it is composed of specialist species, each of which is adapted to a particular environmental condition and can substitute for one another under different conditions. We would expect to see the same taxa appearing in different environmental conditions in a community of generalists, and different taxa appearing in a community of specialists.

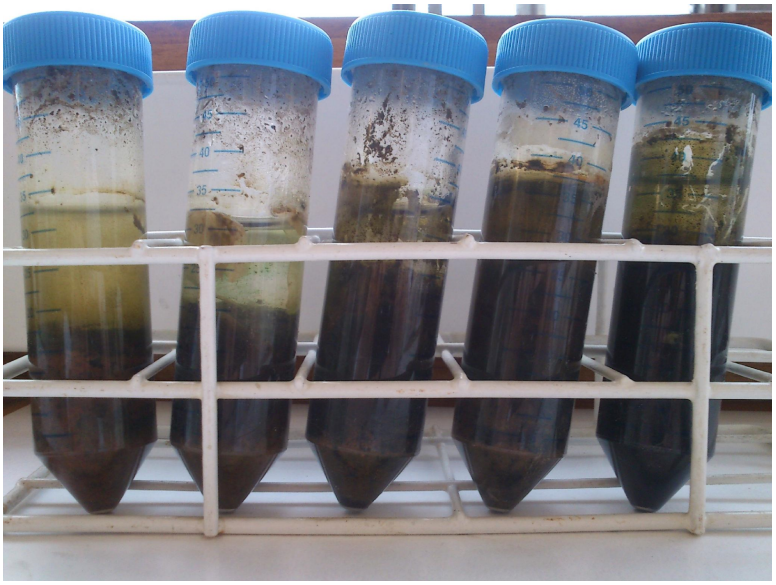


Figure 5.2 *Winogradsky columns at 16 weeks after setup[29]*

To investigate this, a number of model microcosms were set up by Dr Tim Bush [29] to determine the effect of varying cellulose and sulphate concentration on the composition of a community of microbes taken from freshwater pond sediment and allowed to form a nutrient-cycling microcosm, called a Winogradsky column.

A Winogradsky column is a commonly used miniature ecosystem used to study nutrient cycling [10, 135]. To make a Winogradsky column, a sediment sample is taken from a body of water and placed in a small transparent container, overlaid by water from the same source. Additional nutrients are then added, in this case sulphate and cellulose. The column is then incubated for several weeks or months under a light source (Figure 5.2). The nutrients stimulate the growth of bacteria in the sample, and over time a complex nutrient-cycling ecosystem can emerge in the column, driven by energy from the light provided [10, 29].

The ecosystem that develops is these columns is complex. Figure 5.3 shows the key ecosystem processes. Central to the functioning of the community are the carbon and sulphur cycles. In the carbon cycle, degraders break down cellulose to carbon dioxide through a long reaction chain. This carbon dioxide can then be used by oxygenic and anoxygenic phototrophs to create long organic molecules that can be used for energy. These molecules can then be broken down by the degraders. In the sulphur cycle, sulphate is reduced by sulphate reducing bacteria, which use electron donors such as hydrogen or acetate to produce sulphide. Hydrogen and acetate are provided by degraders as a byproduct of degrading biological matter such as cellulose. This sulphide is then oxidised

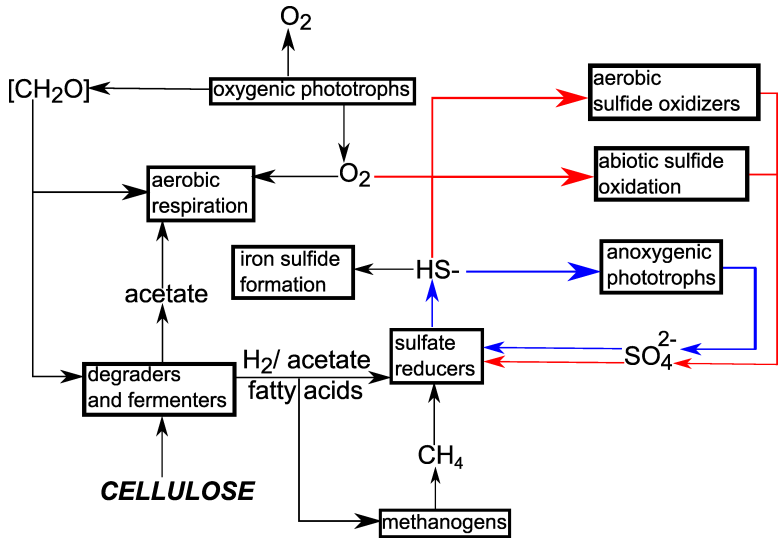


Figure 5.3 *The main processes in our microcosms.*

Diagram of the key ecosystem processes in our microcosms. Processes labelled with a red arrow require oxygen and processes labelled with a blue arrow are anaerobic (reproduced from [29]).

back to sulphate by a number of different paths. It can either be oxidised by anoxygenic phototrophs, oxidised by aerobic sulphide oxidisers, or it can be abiotically oxidised. Alternately it can form iron sulphide, which is then unavailable to microorganisms, as it is insoluble.

In order to understand the communities in the Winogradsky columns more thoroughly, we perturb it by adding different concentrations of nutrients. We can then examine the effect that this has on the microbial community.

In this experiment, samples of sediment and pond water were taken from a local freshwater pond (Blackford Pond). 150 Winogradsky columns were set up with 25g of sediment and 20ml of water in each column. To 132 of these microcosms, cellulose and sulphate were added in varying amounts, with different combinations of cellulose and sulphate concentrations occuring in triplicate[2]. The microcosms were then incubated for 16 weeks and then destructively sampled. From these samples, DNA was extracted and its 16S rRNA gene sequences analysed. One replicate was sequenced from the nutrient concentrations of 0g, 0.19g, 0.28g, 0.38g, 0.47g, 0.56g, 0.75g and 0.94g added cellulose per microcosm, and sulphate concentrations 0 μmol/g, 73 μmol/g and 146 μmol/g, giving a total of 24 sequenced samples. OTUs were then clustered at 97% and rarified.

[2]i.e., three replicates of the same experiment

5.2.1 Assigning functional groups

In our model ecosystem, many essential microbial functions are required to ensure the functioning of the ecosystem, as shown in Figure 5.3. These functions are performed by specific micro-organisms, some of which are known. For known taxa, we can therefore categorise OTUs into functional groups by the function we know them to perform. This allows us to examine the effects of perturbing the system on individual functional groups.

We identify four main functional groups [29]. The full taxonomy of these groups is given in Appendix A:

Oxygenic Phototrophs

This group performs photosynthesis using the light source provided, and therefore provides energy to the rest of the community. They use light energy to reduce CO_2 to organic carbohydrate molecules, using H_2O as an electron donor and releasing O_2 in the process. The O_2 can then be used to oxidise sulphide either abiotically, or by aerobic sulphide oxidising microbes. The organic molecules can be degraded by other microbes to release energy or as a carbon source.

Anoxygenic Phototrophs

This group also photosynthesises, but does not use H_2O as an electron donor and most function anaerobically. They use light energy to oxidise sulphide (HS^-, which functions as the electron donor) to sulphate (SO_4^{2-}). This reaction requires CO_2, which this group compete with the oxygenic phototrophs for.

Degraders

This group obtains energy by degrading the initially added cellulose, and any other complex organic matter in the ecosystem. This is a complex process involving many steps; initially, polymers such such as cellulose are broken down to monomers such as sugars, which are then degraded to organic acids and alcohols. The end products are acetate ($C_2H_3O_2^-$), H_2, which can then be used as electron donors by sulphate reducing bacteria and methanogens, and CO_2, which can be used by oxygenic and anoxygenic phototrophs.

Sulphate Reducers

This group obtains energy by oxidising electron donors provided by the degraders,

reducing sulphate to sulphide in the process. These organisms can be thought of as 'breathing' sulphate rather than oxygen.

Functional groups not included

In Figure 5.3, two functional groups are mentioned that do not feature as part of our analysis.

Aerobic sulphide oxidisers are bacteria that obtain energy by oxidising sulphide using oxygen. No common aerobic sulphide oxidisers were found in our samples.

Methanogens are archaea that obtain energy by reducing carbon dioxide to methane, competing with sulphate reducers for electron donors in the process. Archaea have a divergent 16S rRNA gene from bacteria, and so the bacteria-specific primer used in the PCR reaction do not detect them efficiently.

5.2.2 Identifying functional groups

An OTU table (eg, Figure 5.1) contains a list of samples, a list of taxa, the abundance of each taxon in each sample and the taxonomic information available for each taxon. A table can contain tens of thousands of taxa and sorting them into functional groups is a non-trivial task. A Python script was written for ease of sorting the data. With this script, an OTU table can be input, and OTU tables of individual functional groups, listing either all OTUs or a specified number of high-abundance OTUs, can be output. The script is easy for a non-technical user to edit, to add extra functional groups or to add additional taxa to a functional group. This was important as the aim was to allow biologists in Dr Free's lab to use the script.

A script was written in Python to sort OTU data into functional groups (Figure 5.4). In this script, an OTU object is created for each OTU in the table, an object of the appropriate taxonomic level is created for each instance of a taxonomic level found in the table (i.e., specific classes, families etc) and a Sample object is created for each sample. The Sample object stores the name of the sample and anything else that is relevant, in this case the different values of cellulose and sulphate concentrations. Each OTU object stores the abundance of that OTU in each sample, and is linked to the Sample object using a dictionary. Each OTU

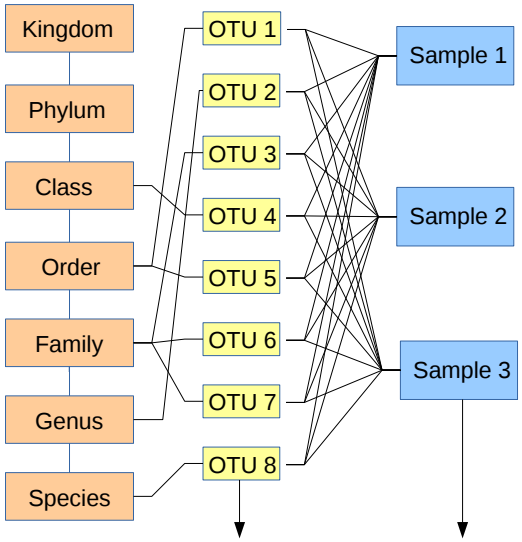


Figure 5.4 *Diagram of data structure used to sort functional groups*

object is also connected to the lowest taxonomic level it belongs to, again using a dictionary. Taxonomic levels are connected together, again using dictionaries.

This structure allows easy partitioning of functional groups; one can simply specify the name of any taxonomic rank and quickly access all OTUs in that taxonomic rank. As dictionaries are used as data structures, the name itself is used as a lookup, rather than a meaningless integer as would be the case if using an array.

With this script, the Shannon diversity and abundance of individual functional groups can be easily computed. In addition, the script contains a function to identify the most abundant individual taxa in a functional group, across all of the samples. The top n most abundant taxa are found, where n is a number specified by the user, realistically between two or ten depending on the diversity. The abundance of these taxa is then found across all samples, making it easy to identify the key taxa in each functional group. In calculating the Shannon diversity, the abundance of each taxon is normalised with respect to the total abundance of its functional group, rather than the total abundance of all taxa.

5.2.3 Analysis and results

To examine the effects of nutrient concentration perturbations on the different functional groups, we first plot the Shannon diversity and abundance in each sample of each functional group (Figure 5.5) and compare this with Figure 5.6,

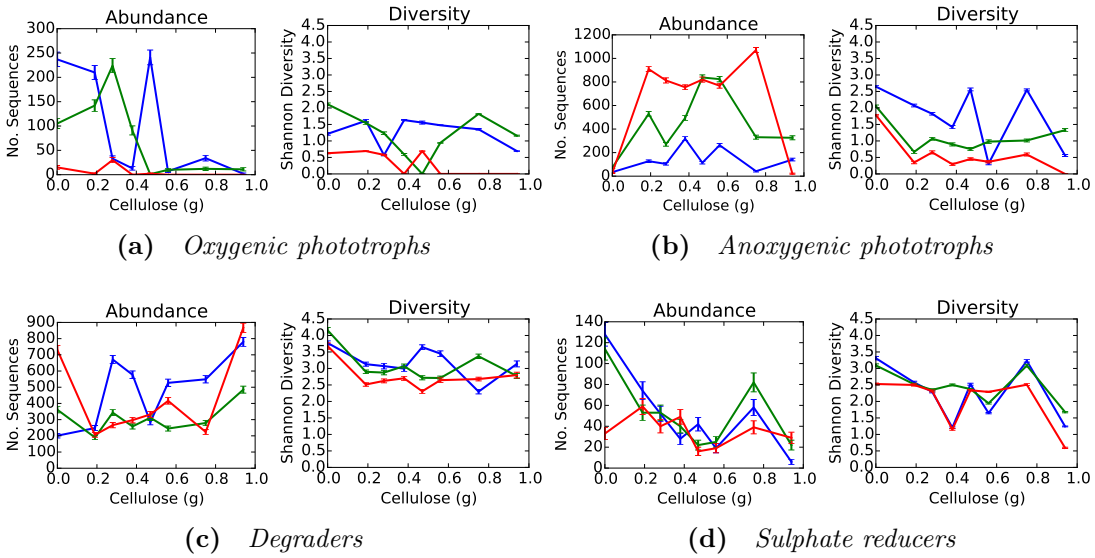


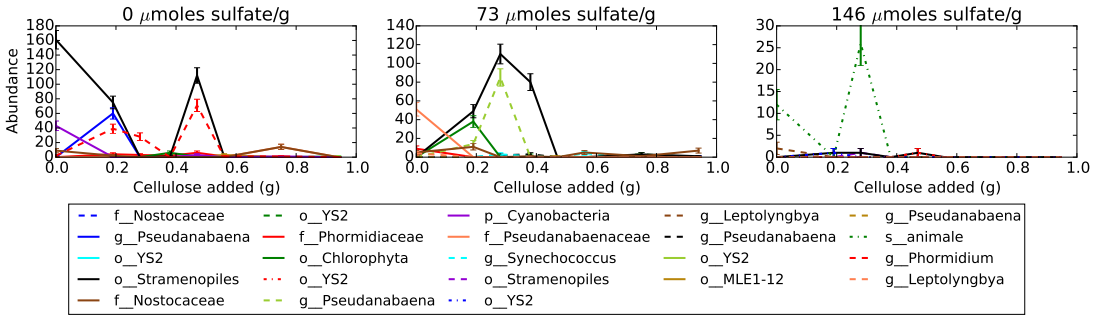
Figure 5.5 *Abundance and Shannon diversity across different functional groups. The Shannon diversity is calculated using taxon abundances normalised with respect to total abundance of each functional group.*

The blue, green and red lines correspond to 0, 73 and 146 µmoles sulphate/g added to the microcosm respectively.

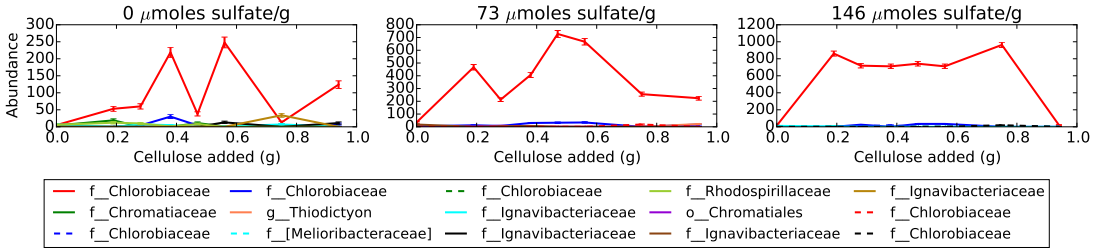
which shows the most abundant taxa across all samples in each functional group. The Shannon diversity of each group is calculated by taking the abundance of each taxon as a proportion of the abundance of the functional group, rather than the abundance of the whole. Calculated in this way, the Shannon diversity of each group can be compared. Taken together, these plots allow us to examine the effect of changing nutrient concentrations on each functional group and establish whether that group is functionally redundant or not.

Oxygenic phototrophs

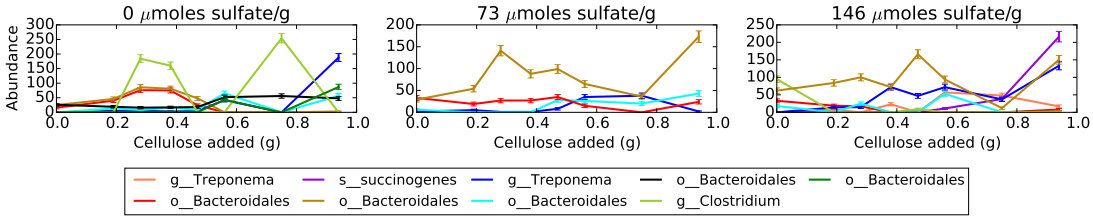
As sulphate increases (blue → green → red lines in Figure 5.5), the oxygenic phototrophs suffer a collapse in both Shannon diversity and abundance (Figure 5.5a). This is likely because oxygenic phototrophs must compete with anoxygenic phototrophs for CO_2, and in a highly sulphidic environment favourable to anoxygenic phototrophs competition will be fierce. In addition, sulphide inhibits oxygenic photosynthesis in many strains of *Cyanobacteria* [106]. However, even at the highest sulphide concentration, some taxa survive (Figure 5.6a). It is likely that those taxa that survive in high sulphate conditions will be those that



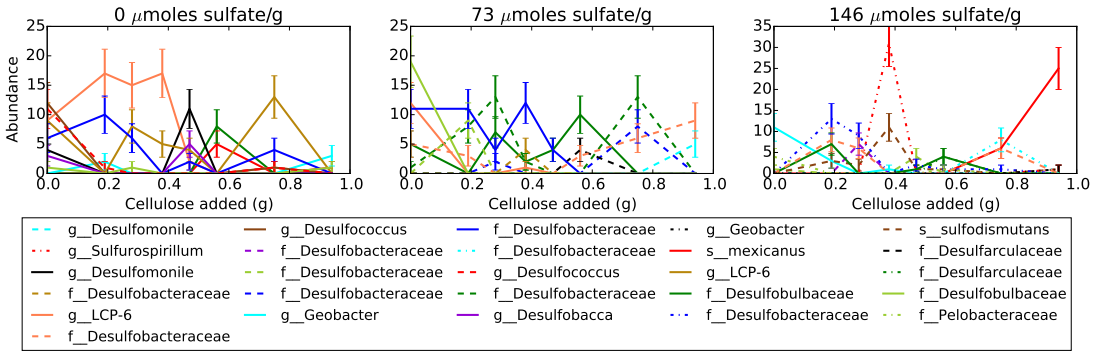
(a) *Oxygenic phototrophs*



(b) *Anoxygenic phototrophs*



(c) *Degraders*



(d) *Sulphate reducers*

Figure 5.6 *Abundant taxa in different functional groups*

The abundance of every taxon that is either the most or second most abundant in one sample across all samples.

can resist sulphide toxicity. In the 146 μmol sulphate/g condition, the only taxon to reach an appreciable abundance (species *animale*) is part of the genus *Phormidium*, which is known to be able to photosynthesise aerobically in the presence of sulphide [35]. *Phormidium animale* is known to survive in a wide range of extreme environmental conditions [141], which suggests that it may be viewed as a generalist among oxygenic phototrophs. That there is no appreciable abundance of *P. animale* in lower-sulphate environments supports this view. In low-sulphate environments, we hypothesise that *P. animale* may suffer a fitness cost by investing in the ability to survive in high-sulphate environments such that it cannot compete with more specialised taxa. However, when the environment becomes more challenging due to high sulphide, it is the only one which survives.

Anoxygenic phototrophs

As the sulphate concentration added to the microcosms increases, so does the abundance of anoxygenic phototrophs (Figure 5.5b). This is to be expected, as a high concentration of added sulphate will result in a high concentration of sulphide, and sulphide is an essential nutrient for the anoxygenic phototrophs (see Section 5.2.1). However, this increase in abundance comes with a dramatic drop in diversity. Examining Figure 5.6b, we find that a single taxon from the family *Chlorobiaceae* dominates at all three sulphate concentrations, but especially so at higher sulphate concentrations (with the exception of the highest sulphate, highest cellulose concentration, which is anomalous). This is likely because of the artificial nature of our experiment. In the natural environment of a pond, different members of the family *Chlorobiaceae* specialise on different light intensities at different depths [110]. In our microcosms, this range of light intensities is removed. As all anoxygenic phototrophs are competing for exactly the same resources of CO_2 and HS^-, the homogeneous light intensity places them all in an effectively homogeneous environment. In this type of environment, the taxon most specialised to that environment wins, as we see here.

Degraders

The abundance and diversity of degraders does not change a great deal with changes in additional sulphate or cellulose (Figure 5.5c). We also note from Figure 5.6c that many of the same taxa appear in our microcosms in changing

environmental conditions. Two taxa of Order *Bacteroidales*, from the Phylum *Bacteroidetes* (brown and red lines, Figure 5.6c) appear across the full range of sulphate concentrations and the full range of cellulose concentrations, as does the taxon of Order *Treponema*, from the Phylum *Spirochaetes* (blue line, Figure 5.6c). This suggests that these taxa are generalists, and can survive in many different environments, perhaps by using alternative metabolic pathways. However, we also note that not all successful taxa appear to be generalists - *Fibrobacter succinogenes*, from the Phylum *Fibrobacteres* is the highest abundance taxon in the highest sulphate and highest cellulose concentrations (purple line) and does not appear in high abundance in any other conditions. This species uses a unique method of cellulose degradation and is thought to be a specialist [152], in agreement with this observation.

Sulphate reducers

Similarly to the degraders, the abundance and diversity of sulphate reducers stays roughly constant across the range of environmental conditions (Figure 5.5d). However, unlike the degraders we also see that the most abundant taxa in each environment change considerably (Figure 5.6d). All but four of the taxa that appear in environments with the highest sulphate concentration do not appear at all in lower sulphate environments. Indeed, in many conditions we observe only one or two with any substantial abundance and some high-abundance taxa appear only once across all the environmental conditions (*Sulfurospirillum*, red dotted line and *Desulfomonile*, black line). This is reflected in a low diversity across all conditions (Figure 5.5d). We tentatively speculate that this is due to sulphate reducers being a community of specialists, each of which functions well in a specific environmental condition, but cannot function well once that environmental condition is deviated from. It is likely that sulphate reducers have adapted to occupy micro-niches in the natural environment, which would allow them to coexist [40]. However, in our microcosms the environmental heterogeneity required for this coexistence is not available. Although we do not see coexistence in individual microcosms, the group as a whole demonstrates functional redundancy, as despite these large changes in taxa composition, the overall diversity and abundance is relatively unchanged, allowing the group to continue to function as normal with regard to the rest of the ecosystem. It is however important to note that only one replicate from each nutrient set has been sequenced and so we cannot be sure that the variation seen is not due to

simple stochastic variation. Sequencing each of the triplicate microcosms could provide a more robust conclusion.

5.2.4 Summary of nutrient varying experiment

In this section, we have used functional group analysis to explore the effects of varying nutrient composition on a microbial community. We find that the degrader and sulphate reducer functional groups are stable to changing environmental conditions, with relatively unchanged values of diversity and abundance, while the oxygenic and anoxygenic phototroph communities change dramatically in diversity and abundance as we vary the environmental conditions.

We speculate that the oxygenic phototroph, anoxygenic phototroph and sulphate reducer functional groups are composed mostly of specialists. This is based on a number of observations. Firstly, the diversity in all of these groups is relatively low across all environmental conditions, and very low in others. In the natural environment, a community of specialists may have a high diversity, as there may be many environmental niches that may be filled by specialists that coexist alongside each other. However, in our small microcosms, we speculate that there is not enough space for the environmental heterogeneity required and each microcosm is effectively a homogeneous environment.

In the case of the oxygenic phototrophs, we observe few taxa (in low or medium sulphate concentrations) or no taxa at all (in high sulphate concentrations). In the high sulphate case, all but one taxon is unable to survive in the highly sulphidic environment. We speculate that this taxon (*Phormidium animale*) is a generalist in the oxygenic phototroph group as it is known to still function in a range of extreme environmental conditions [141], as well as in our high-sulphate microcosms.

In the anoxygenic phototroph community, one single taxon in the family *Chlorobiaceae* dominates across all microcosms. We speculate that this is because niche partitioning in *Chlorobiaceae* in the natural environment depends on different taxa specialising on different light intensities. In all of our microcosms, only a single light intensity is available, and so the lucky taxon which is specialised to it succeeds exceptionally. We speculate that if a number of microcosms were set up with different available light intensities, we might see a different *Chlorobiaceae* taxon established in each microcosm.

In the sulphate reducer community, we see a wide range of taxa in different environmental conditions, but only a small number in any individual condition. This suggests that in nature, sulphate reducers form a diverse and specialised community, with each taxon specialised to their own environmental niche. We speculate that sulphate reducers niche partition across different chemical concentrations, and therefore prosper in different microcosms (in contrast with the anoxygenic phototrophs, where we provide only one instance of light intensity, so observe the same taxon in each microcosm).

By comparison, we speculate that the degrader community is made up mostly of generalists. This community has high diversity across all environmental conditions (Figure 5.5c) and we see that taxa both maintain high diversities across a range of environmental conditions, and coexist with one another (Figure 5.6c).

5.3 Network analysis on multiple replicate microcosms

We now turn to a second microcosm experiment, set up by Dr Euyln Pagaling in the Andrew Free lab. The aim of this experiment was to investigate factors that influence microbial community composition - in particular, whether microbial community assembly is deterministic or stochastic at small system sizes [122].

Previous work [121] had found that experiments with microcosms in triplicate that were seeded from the same source developed different microbial communities. This experiment aimed to investigate this in more detail, using many more replicates.

In this study, samples of pond sediment were taken, and used to set up 100 identical Winogradsky columns. In the previous work [121], it was found that development of microcosms was complete after incubation for 16 weeks, with very little change in community composition after this. For this reason, half of the microcosms were incubated for 8 weeks and then destructively sampled and the other half were left for 16 weeks and then destructively sampled. This should give information about communities during the early and late stages of development.

The fact that many microcosms are set up in identical conditions made this experiment an excellent candidate for finding correlations between taxon abundances, which might indicate interactions between the taxa. If a correlation is found between two taxa in microcosms with different environmental conditions this might be simply because both taxa are independently adapted to the same set of conditions. However, if all microcosms are identical replicates, the co-occurrence of two taxa in some of them, but not others, should imply that they are having a positive effect on one another, i.e. interacting mutualistically.

We can then compare the results of this analysis with the experiment detailed in the previous section to determine whether generalist or specialist species tend to co-occur.

5.3.1 Network Analysis

A 'network' is a system in which units are connected, in some way, to one another. Units are referred to as 'nodes' and the connections between them as 'edges'. In an ecosystem network, nodes denote a particular type of organism, and edges between nodes indicate some interaction between those organisms such that the abundance of one organism affects in some way the abundance of the organism it is connected to.

Network analysis can play an important role in understanding the role that microbial organisms play in an ecosystem. In particular, if links can be found between one taxon and many other taxa, that taxon is a 'hub' (node with many edges) in the network and it can therefore be assumed that its presence is important to the ecosystem function as a whole.

To determine the links between taxa using 16S rRNA sequence data, the co-occurrence patterns between taxa must be examined across multiple independent samples. If two taxa have high abundances in the same samples, and low abundances in the same samples, they are positively correlated, and may have an underlying mutualistic interaction. If one taxon has high abundances in the samples in which the other taxon has low abundances, they are negatively correlated, and may have an underlying antagonistic interaction.

The large number of replicate microcosms in this experiment made it a prime candidate for network analysis, as it is necessary to have a large number of samples for correlations to be detected with statistical significance.

The Spearman rank correlation coefficient

A correlation coefficient is a measure of how correlated two random variables are across multiple datasets. In the context of finding correlations between bacterial taxon abundances, each taxon is considered to be a random variable that may have a different value in each sample.

Numerous correlation coefficients can be used, but the standard to use for this type of network analysis is the Spearman rank correlation coefficient (ρ). ρ falls between 1 and -1, where 1 indicates a perfect positive correlation, -1 indicates a perfect negative correlation, and 0 indicates no correlation.

Sample	Variable A	Variable B	Rank A	Rank B	d	d^2
1	4	2	5	3.5	1.5	2.25
2	5	1	4	5	1	1
3	12	2	3	3.5	0.5	0.25
4	13	13	2	1	1	1
5	100	10	1	2	1	1

Figure 5.7 *Illustration of the calculation of Spearman's ρ.*

Spearman's ρ is a rank based correlation coefficient. This means that rather than using the actual abundance of the random variables to calculate the coefficients, the samples are ranked in order of maximum to minimum abundance of a particular taxon in each sample. The coefficient is then calculated using the difference in rank of each sample. Table 5.7 shows an illustration of the process. Here variables A and B represent the abundances of two taxa in samples 1,2,3,4 and 5. In order to calculate Spearman's ρ between the two variables, the samples are ranked for each variable in order of abundance in that sample (The sample with the highest abundance of A is given a rank of 1 for A, and so on). If more than one samples have the same value for the variable, as in the case of samples 1 and 3 for variable B, the rank is the same for all of them. The difference, d, in ranks between the two variables is found for each sample and then squared.

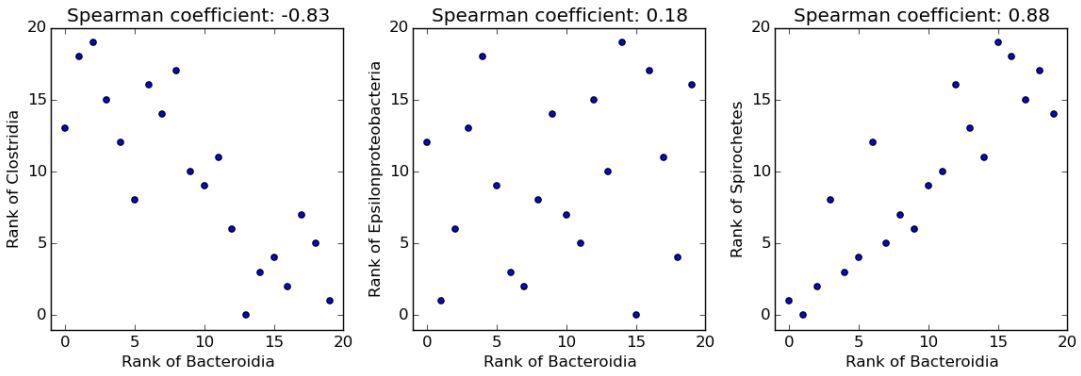
$$\text{Spearman } \rho = 1 - \frac{6 \sum d^2}{n(n^2 - 1)} \tag{5.1}$$

d^2 can then be used in Equation 5.1, where n is the number of samples, to find Spearman's ρ. In this illustration, ρ is equal to 0.72, with a p-value of 0.17, calculated using the Scipy stats package Spearmanr. This is a high positive correlation, but the p-value is not significant. For a significant p-value, more samples than five must generally be used.

5.3.2 Methods

The methods described in this section follow Barberán et al [16].

Using sequence data from our microcosms, OTUs were clustered at 90% similarity to produce an OTU table with 2630 OTUs. This is a lower threshold of similarity to the standard of 97%, but results in fewer OTUs and thus less complexity. OTUs



(a) **(b)** **(c)**

Figure 5.8 *Illustration of ranks using data from the multiple replicates experiment*

that appeared with a nonzero abundance in at least half of the samples were then selected to make a network. This resulted in 448 OTUs, preserving 98.6% of the total abundance. This pruning step greatly reduces the network complexity, while preserving the vast majority of the abundance. A different requisite number of samples could have been used if necessary to reduce complexity further, or to include more taxa if too much abundance was being lost. Spearman rank correlation coefficients were calculated between these OTUs, using the Scipy Stats package Spearmanr in Python.

An illustration of ranks between high-abundance taxa is shown in Figure 5.8, using real data from the microcosms. Here, the rank abundance of the Class *Bacteroidia* is compared with the Classes *Clostridia*, *Epsilonproteobacteria* and *Spirochetes*. Each dot shows a sample, with the rank of each abundance plotted. *Bacteroidia* is negatively correlated with *Clostridia*, as samples in which *Bacteroidia* has a high abundance are the same samples in which *Clostridia* has a low abundance. *Bacteroidia* is uncorrelated with *Epsilonproteobacteria*, and positively correlated with *Spirochetes*.

A link was considered valid if the Spearman's coefficient ρ was both >0.55 and with a p value of <0.01. The 'weight' of the link was assumed to be equal to ρ. This resulted in a network containing 445 OTUs, or 98.5% of the total abundance. Using the Python package Networkx, a network was then constructed, where nodes were taxa and edges were correlations between taxa.

An algorithm to visualise and identify clusters

We now present a new method used to visualise clusters in the network of taxa. The network is shown in an adjacency plot, where the positioning of each node is chosen to maximise clustering. This method is appropriate for non-directed networks with symmetric links, as the adjacency plot must be symmetric along the diagonal.

To carry out this algorithm, a measure of 'disorder' of the network was used, which was defined as the weight of each positive edge ($\rho_{i,j}$, being the Spearman correlation coefficiant ρ between two taxa i and j), multiplied by its distance to the diagonal, summed over all positive edges, summed over all rows (Equation 5.2).

$$\text{Disorder} = \sum_{i}^{N} \sum_{j}^{N} \text{abs}(i-j)\max\{\rho_{i,j}, 0\} \tag{5.2}$$

The disorder is therefore at a minimum if high positive weight edges are positioned close to the diagonal, which therefore means that nodes with strong positive links with one another are positioned adjacent to one another. This results in clusters of nodes being clearly visible. To attain this optimum configuration, nodes are first placed in a random order. Nodes are then selected at random and then their positions were swapped if the act of swapping would result in the disorder across the network decreasing. We are not aware of this method having been used before.

5.3.3 Results: Identifying clusters for 16 week microcosms

NMDS analysis was carried out by Pagaling et al [122] to determine the similarity of microcosm communities to one another. NMDS (non-metric multidimensional scaling) is a way of visualising the similarity of multiple datasets [7]. In NMDS analysis, the variability of each variable (in this case a particular taxon) is evaluated across multiple datasets, with each variable being treated as an independent 'axis'. To visualise dissimilarity on a two-dimensional plot, the two axes with the most variability are picked and datasets are placed according to the value for these two variables.

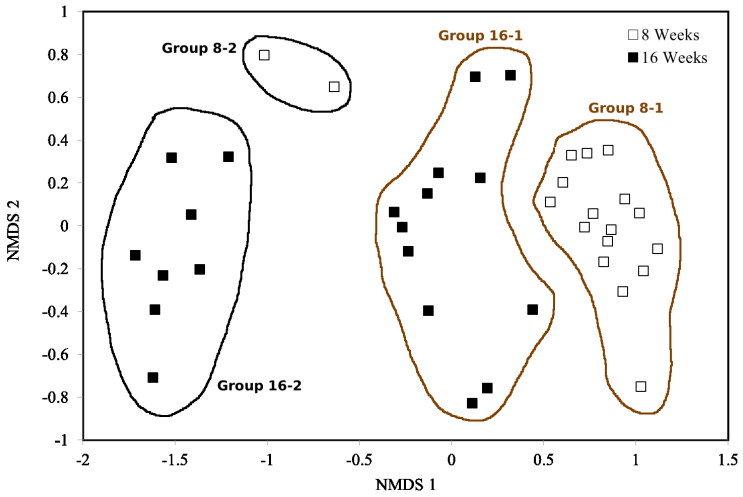


Figure 5.9 *NMDS plot of 8 and 16 week microcosms*

Each dataset is a microcosm sample, and the distance between each sample indicates the similarity of the microbial community found in each sample. Brown and black circles surround groups 1 and 2, respectively (adapted from Pagaling et al [122]).

If two datasets are similar, they will appear on an NMDS plot as close to each other on a non-dimensional scale. This analysis showed that the 16 week microcosms clearly separate into two groups of roughly equal size, and the 8 week microcosms into two groups, one much larger than the other. This suggests that separate microbial communities are found in the two groups of microcosms. We label these groups as 16-1 and 16-2 for the 16 week microcosms, and 8-1 and 8-2 for the 8 week microcosms, where 16-1 and 8-1 appear close together, as do 16-2 and 8-2 (Figure 5.9). The aim was then to identify the taxa that make up these communities, and to determine the community structure through network analysis.

We then use the clustering algorithm outlined in the previous section to visualise the network. We can see clearly that two subnetworks of taxa emerge from the data (Figure 5.10), which are negatively correlated with one another. The taxa within these subnetworks were assigned into categories Cluster 1 (in purple) and Cluster 2 (in green), with the boundaries between clusters in Figure 5.10 drawn by eye. This results in a list of OTUs that belong to one cluster or the other, with a very small number of OTUs that belong to neither. The abundances of this set of OTUs can then be examined in individual microcosms, to determine if a particular microcosm contains only OTUs from one cluster, or if taxa in both of the two subnetworks coexist in the same microcosm (Figure 5.11). We find that

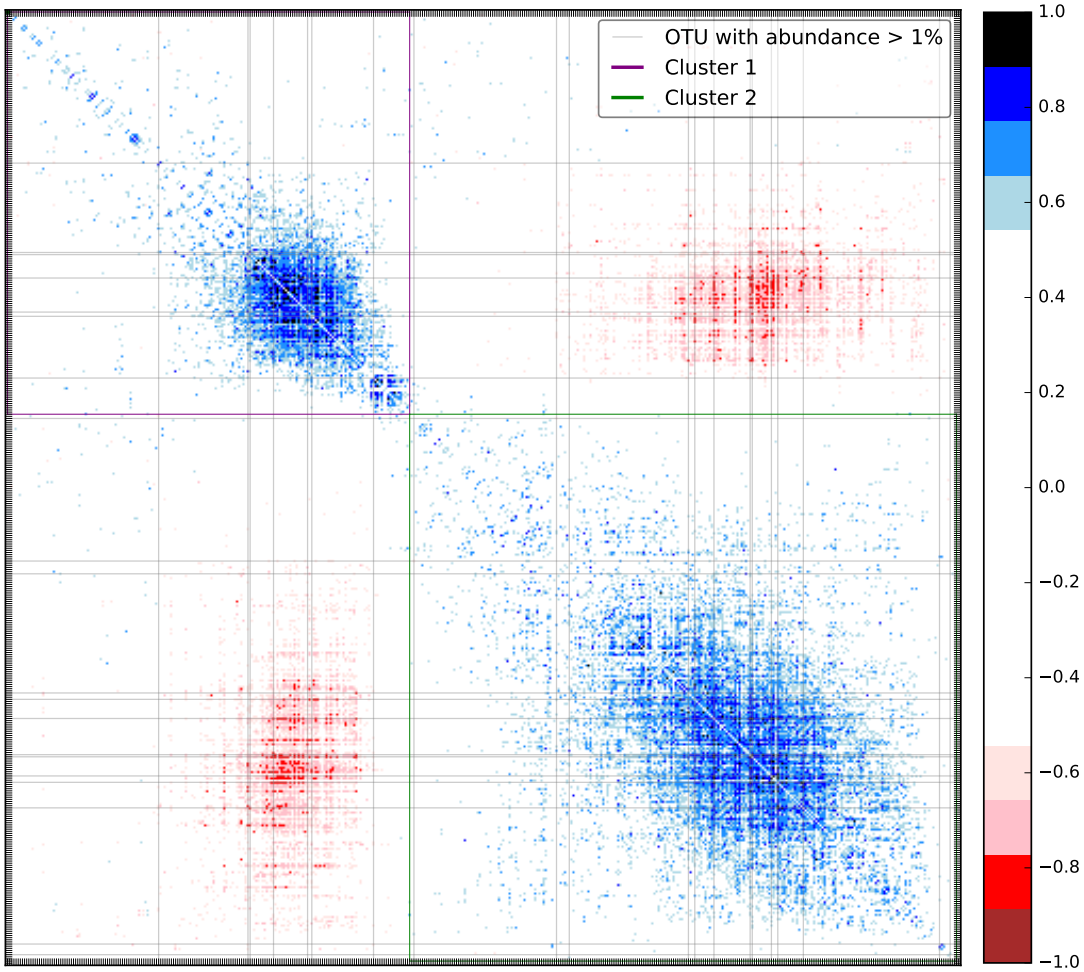
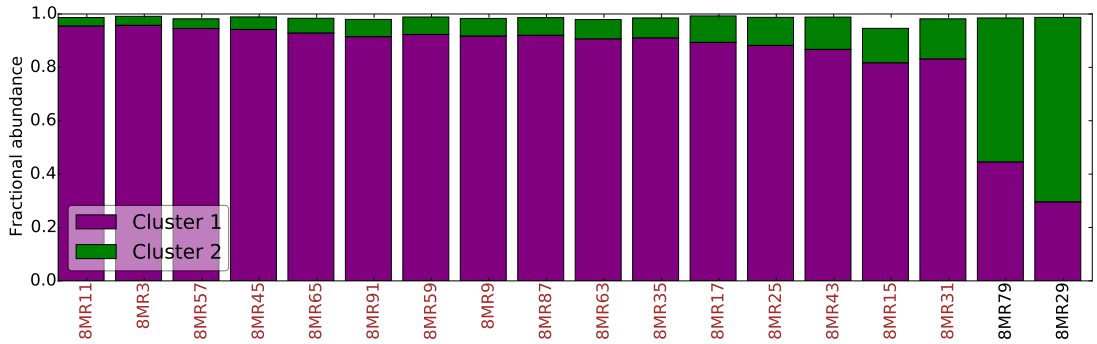
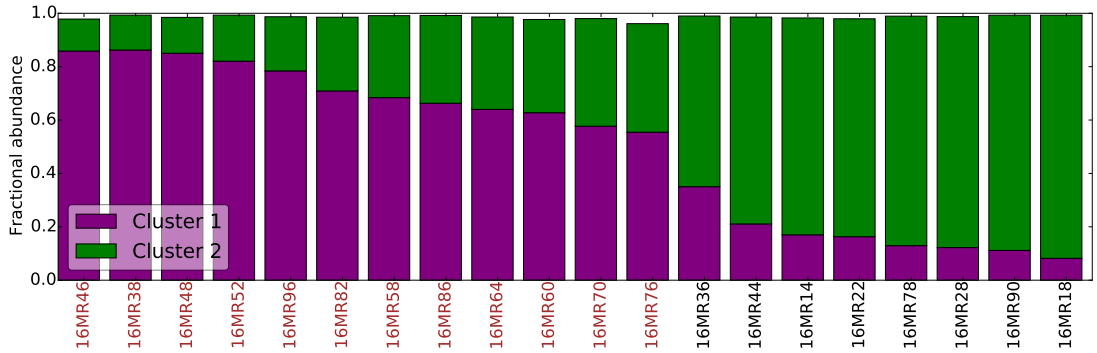


Figure 5.10 *Network for 16 week samples*

The full network for samples at 16 weeks. Positive correlations are shown with blue squares and negative correlations with red squares (see colour bar to the side). There are clearly two large positive clusters and the centre of each cluster is anticorrelated with the centre of the other cluster, shown by the red clusters. OTUs with a mean abundance of more than 1% of the total abundance are shown with grey lines; we see from this that high abundance OTUs are found throughout the network. The network has been separated into two clusters 1 and 2, shown by the coloured squares drawn around them in purple and green, respectively. These have been drawn by eye.



(a) *8 week samples*



(b) *16 week samples*

Figure 5.11 *Abundances of each group in each sample at 8 and 16 weeks.*

The total abundances of each sample at 8 or 16 weeks, divided by the taxa found in clusters 1 (purple) and 2 (green) (Figure 5.10). The sample names in groups 16-1 and 8-1 are coloured brown, those in 16-2 and 8-2 are coloured black.

all microcosms contain a majority of OTUs from either cluster 1 or cluster 2, with a clear divide between microcosms having majoritively one or the other. We also find that this matches up exactly with the groups defined from the NMDS plot (Figure 5.9), in that all microcosms from groups 16-1 or 8-1 contain majoritively OTUs from cluster 1, and similar for 2.

5.3.4 Examining the communities at 16 weeks

To easily see identify key OTUs and their functionality, we can redraw Figure 5.10, only including high-abundance OTUs (Figure 5.12). In this plot, OTUs are colour-coded according to their functionality. We can see a number of interesting things from this plot.

Firstly, we see that the key degraders (purple labelling) in Cluster 1 all belong to

the phylum *Firmicutes*, while the degraders in Cluster 2 almost all belong to the phyla *Bacteroidetes* or *Spirochaetes*. *Bacteroidetes* and *Spirochaetes* are highly positively correlated with one another, and negatively correlated with *Firmicutes*.

The positive correlation in abundance between *Bacteroidetes* and *Spirochaetes* taxa may be because *Spirochaetes* is advantaged in the presence of *Bacteroidetes*. *Bacteroidetes* is known to be cellulolytic (breaks down cellulose) [87] and *Spirochaetes* is known to associate with cellulolytic bacteria [66], as they obtain energy from soluble sugars released from cellulose by cellulolytic bacteria.

Interestingly, a split between *Bacteroidetes* and *Firmicutes* communities has also been seen in previous work on gut microbial communities. Studies have shown that the gut microbiota of obese mice [158, 159] and monozygotic twin humans [160] contain elevated populations of *Firmicutes*, while those of their lean counterparts contain elevated populations of *Bacteroidetes*. In addition, transplantation of microbiota from obese subjects to lean subjects resulted in lean subjects depositing more fat than those following transplantation from a lean subject [159].

Secondly, in Figure 5.12, we see that there are a great deal more taxa whose functions are unknown (green labelling) in Cluster 1 than in Cluster 2. Of these, the taxa belonging to the class *Anaerolineae* are particularly intriguing, being highly positively correlated with the key *Firmicutes* taxa in Cluster 1, and highly negatively correlated with the key taxa *Bacteroidetes* and *Spirochaetes* in Cluster 2. This suggests that *Anaerolineae* is important in the community functioning of Cluster 1. The function of *Anaerolineae* is relatively unknown; however it is known that a cultivated species within this class (*Anaerolinea thermolimosa*) can utilise sugars as a substrate for growth [176]. As *Spirochaetes* are also known to use substrates for growth excreted by other bacteria [66], it may be reasonable to speculate that *Anaerolineae* uses cellobiose and other intermediate cellulose degradation compounds that are excreted by *Firmicutes* as substrates, and that they have a symbiotic relationship with *Firmicutes* analogous to that between *Bacteroidetes* and *Spirochaetes*. Therefore, this work may be used to further understanding of the function of *Anaerolineae* and its role.

In addition, we note that the unassigned taxon with class *Verruco-5* in Cluster 2 is high-abundance, very strongly positively correlated with key taxa in Cluster 2 and negatively correlated with *Firmicutes* in Cluster 1, in addition to being very strongly negatively correlated with *Anaerolineae*. *Verrucomicrobia* have been

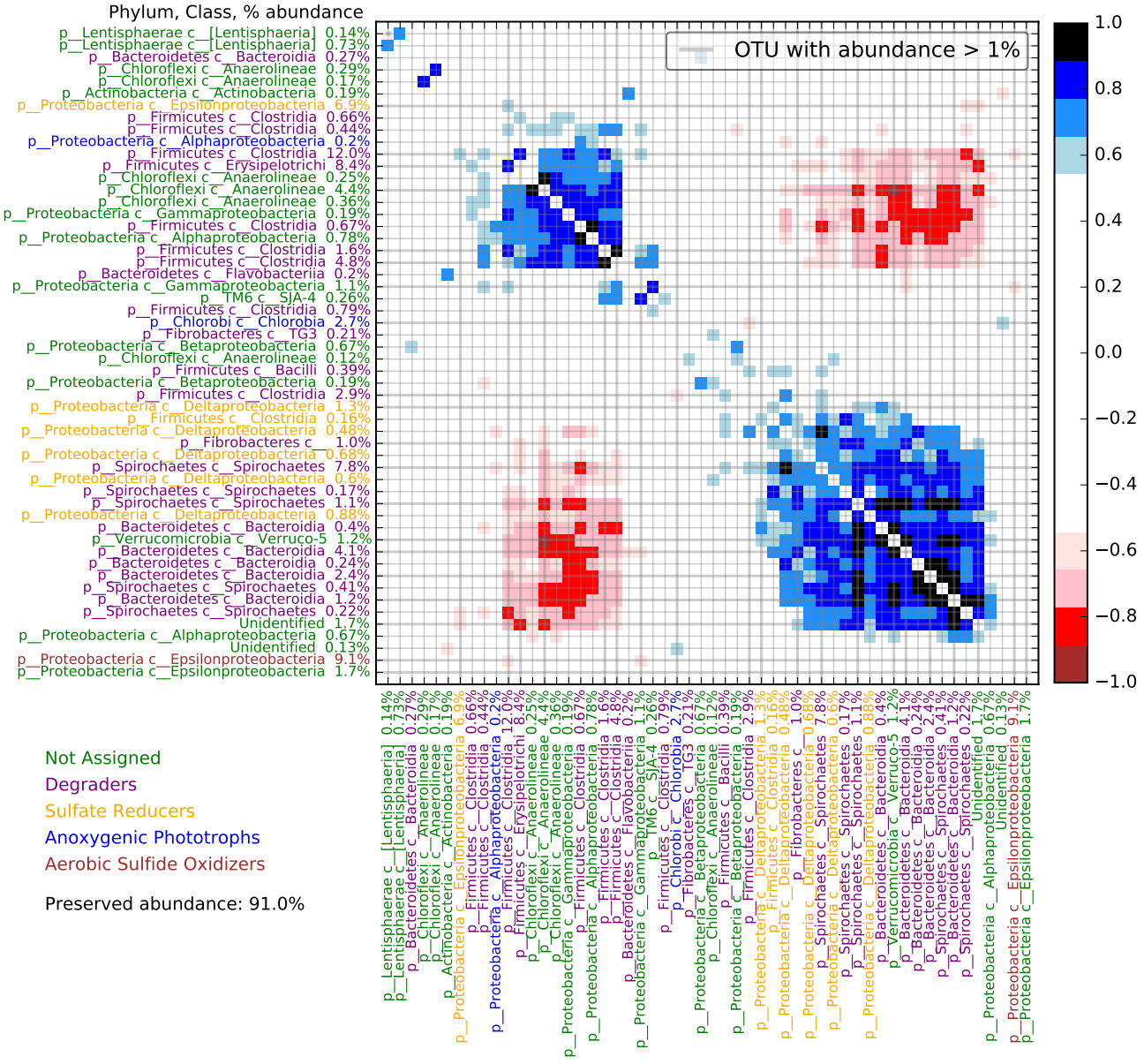
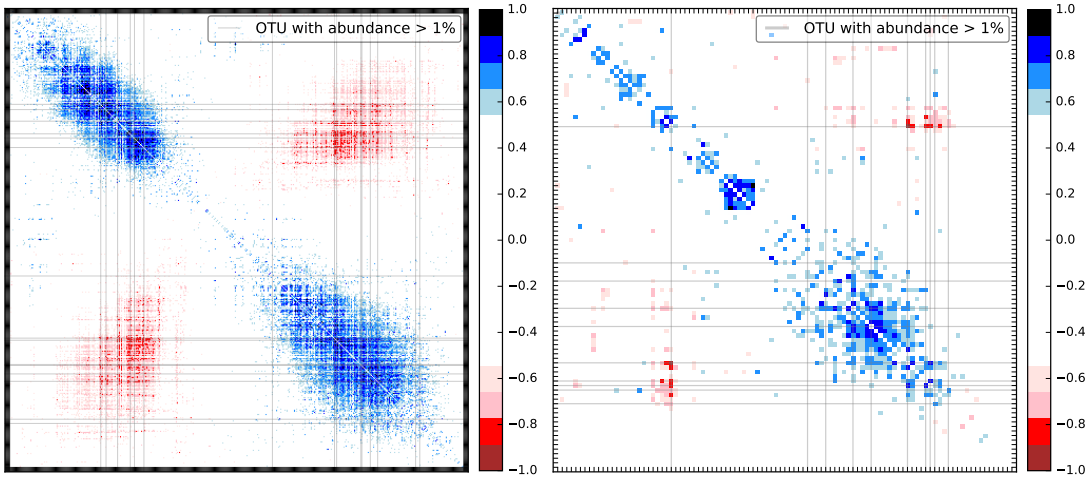


Figure 5.12 *Communities at 16 weeks, low abundance OTUs removed*

Communities at 16 weeks, visualised in Matplotlib using a script written in Python. The same OTU order as in Figure 5.10 is used, but only including OTUs with either mean abundance across all samples >0.1%, or a maximum abundance across all samples >1%. This preserves the abundance of taxa shown as 91% of the total abundance. Taxa are labelled at phylum and class level and colour-coded by their functional group (legend in bottom left corner).



(a) *Degraders* **(b)** *Sulphate reducers*

Figure 5.13 *Networks of degraders and sulphate reducers*

found to contain genes which are associated with cellulose degradation [168], and further investigation into the function of *Verrucomicrobia* would appear to be a prime area for further investigation to establish whether it can be classified as a degrader too.

5.3.5 Network analysis within functional groups

We can also use the same approach to examine correlation networks between taxa within functional groups (Figure 5.13). This analysis reveals that there are clearly strong subcommunities within the degrader group. In comparison, correlations between sulphate reducers are weaker, and the sulphate reducer taxa do not appear to split two clusters.

5.3.6 Summary of multiple replicate microcosms

The network analysis of the data from this multiple replicates experiment uncovers a number of interesting facts. Firstly, the microbial community divides into two sub-communities. The hub (highly connected) nodes in each sub-community, which also tend to be high-abundance, tend to be significantly negatively correlated with the hub nodes in the other sub-community. We therefore speculate that the two communities tend to compete with one another, such that they cannot stably co-exist.

Previous work has shown that by 16 weeks, the community in these microcosms is stable and not subject to further change. Our results support this by showing that there is a clear gap between the community composition of a group A microcosm and the composition of a group B microcosm (Figure 5.11b), meaning that each microcosm has a single sub-community dominating in each. Examining just the degrader community (Figure 5.13a), this gap is even clearer, suggesting that a network of degrader sub-communities form the base of the overall sub-communities, with taxa from other groups interacting mostly with degraders and not with one another. This split appears to lead to group selection. If one group of taxa are adapted to work with each other, and another group are also adapted to work with each other, putting the two groups together in a confined environment will result in one group winning, which we see in our results.

However, this selection does not appear to be deterministic. All microcosms are set up and maintained under identical conditions. The large deviations that we see in the community makeup must therefore come from stochastic variations in abundances leading to nonlinear feedbacks, which eventually leads to a split in the community. In this split, a community either becomes *Firmicutes* or *Bacteroidetes* rich. This split is particularly interesting because of the link between *Firmicutes*-rich gut microbiota and obesity, as it may suggest that an individual may become obese simply by chance.

5.4 Conclusions

In the first section (5.2), we found that degraders are the most generalist community in the functional groups we have studied. In the second section (5.3) we find that the degraders are the most highly correlated group, suggesting that there are many mutualistic relationships between different degrader taxa. This may explain how so many taxa performing the same functional task can coexist in the same constrained, homogeneous environment of our microcosms (shown by the high diversity of degraders in Figure 5.5c). Different taxa may perform different jobs in the overall task of degrading a molecule, allowing niche partitioning in this way rather than the niche partitioning of consuming different resources or living in different environmental conditions. It would be an interesting area of further study to investigate if generalist macroscopic species are more likely to form mutualistic interactions for survival than specialists which depend only on a specific type of environment.

Chapter 6

Conclusions

In this thesis, we have explored a number of ways in which species in the same ecological community may interact.

In our theoretical models, we have shown in our theory of resource spectrum engineering that species on the same trophic level competing for similar resources may exert indirect commensal or mutualistic forces on one another; for example, the presence of extreme and intermediate specialists results in more mild specialists. In this mechanism, one species may benefit another by suppression of a third competitor shared by both [169]. The presence of this interaction has been shown experimentally in small numbers of species in field experiments with old-field [107], riparian [91] and salt marsh [30] plant communities, as well as demonstrated theoretically using Lotka-Volterra models [151]. However, this has not previously been linked to specialist-generalist coexistence. Our theoretical work indicates that, as a result of resource spectrum engineering, the presence of generalists benefits other generalists, while the presence of specialists benefits other specialists in an ecological community. Indeed, a study of bird communities by Juillard et al [82] find that specialist bird species are more likely to be found in habitats with other specialists and vice versa for generalists, supporting this view.

Our theoretical work considers uses static environments; i.e. those that do not change in time. As a result, our model demonstrates only a community-wide dynamical shift from widespread generalism to widespread specialism, as the main mechanism favouring generalists is not included. However, it is reasonable to assume that including temporal variations in the environment may result

in our communities shifting from specialist-dominated to generalist-dominated. Recently, the specialist-generalist balance in many communities in the natural environment has been shown to be shifting towards generalism, probably as a result of habitat disturbances caused by humans [33]. This has been seen in communities of plants [134], birds [41, 164], mammals [54] and fish [112]. The resource spectrum engineering mechanism discussed in this thesis suggests a means by which this process could be accelerated; the loss of small numbers of specialist species in disturbed habitats could allow more generalist populations to grow larger, causing them to then displace specialists in undisturbed habitats. Further work involving modelling environments with temporal disturbance could therefore be very interesting.

Turning to spatially heterogeneous simulations in which dispersal is limited in our theoretical model, we found that a community of generalists favoured long dispersal ranges, while communities of specialists favoured short dispersal ranges. In addition, spatially uncorrelated environments favoured communities consisting of either generalists or highly specialised species, depending on the dispersal range, while more correlated environments favoured more mild and intermediate specialists. We find that, paradoxically, less correlated environments favour more specialised species for intermediate dispersal ranges.

This provides an example of a difference between the strategies of intermediate and extreme specialists. When dispersal is limited, an intermediate specialist may colonise a local area in which resources are similar, but may not be able to survive if those resources are scattered in space. By contrast, an extreme specialist may be able to use scattered amenable resources even if its propagules do not disperse very far, because the few that disperse far enough will be able to displace less specialised competitors. Intermediate specialists, similar to generalists, require a larger amount of accessible amenable habitat than extreme specialists do.

It is important to note once again that our model does not include temporal variation in the environment, which almost certainly would encourage short-range dispersal in specialists. If disturbance was included in the environment, specialists would be forced to disperse larger distances in order to find new suitable habitat. In support of this idea, a study by Entling et al [50] of aerial dispersal in spiders concluded that the survival of specialist species with limited dispersal ranges requires a stable habitat.

In our final chapter we analyse sequence data from a series of microbial microcosm

experiments. By separating these micro-organisms into different functional groups, we find that some groups appear to behave as generalists, being able to thrive in a variety of environmental conditions, while some are more specialist, only being able to survive in a limited range of conditions.

In the network analysis performed on multiple replicate microcosm experiments, we found many co-occurrence patterns between different taxa. We found that many taxa have positive associations with one another, suggesting mutualistic or commensal interactions. The microbial network separated into two distinct community clusters. The small number of negative correlations that were found, which suggested a competitive or amensal interaction between species, were found between members of the two communities, suggesting that the communities competed with one another as a group. This two-community effect was more pronounced when the degrader group was considered on its own, suggesting that degrading micro-organisms have many interactions with one another. This may be because the process of degrading biological matter uses many intermediate steps, so a number of taxa may preferentially work together and evolve to function optimally in conditions caused by another specific taxon.

In this thesis, we have developed a theoretical model ecosystem that we have used to explore the effect of the environment on the specialist-generalist balance, and contributed our theory of resource spectrum engineering to the literature. An obvious first step for further work using this model would be to include an environment that changed over time. This change could be regular and periodic, similar to the changing of the seasons in the natural world, or unpredictable, similar to natural disasters such as forest fires. Another possible extension could be to add another niche axis, so that species could specialise or generalise on more than one. For example, 'habitat type' and 'resources' could be simulated as independent entities. It would be interesting to explore whether successful species could specialise on one axis and generalise on the other.

Experimental testing of the theory of resource spectrum engineering could feasibly be done in the future using micro-organisms. Standard genetic engineering methods mean it is possible to precisely engineer strains of *E. coli* that are able to metabolise different types of sugars. In particular, one could engineer a 'generalist' strain that can metabolise many types of sugars, an intermediate specialist could metabolise a small number of sugars and an extreme specialist could metabolise only one sugar. These strains could then be placed in competition with one another in a chemostat that contains a variety of sugars.

One could start the experiment with many sugars in a small abundance as to favour generalists, but then gradually increase the abundance of specific sugars to favour particular specialist species. If this manipulation caused the generalist to be displaced even on the sugars that had not increased in concentration, the theory would be verified.

In our analysis of experimental microcosms, we have demonstrated that a microbial community may have alternative stable states. It would be very interesting to apply this analysis to other microbial systems to determine whether these states emerge for other communities.

Appendix A

Functional group assignments

This appendix lays out the full taxonomic information of the taxa in our functional groups[29]. Only taxa that were found in either the variable nutrient concentration experiment or multiple replicates experiment are included.

If a taxon is included twice (for example, the genus *Desulfotomaculum* is defined as a sulphate reducer even though it is in the phylum *Firmicutes*, all of which has been defined as a degrader, the functional groups with higher classification takes precendence and is removed from the lower classification functional group, ie *Desulfotomaculum* does not appear as a degrader.

Oxygenic phototrophs

- Phylum *Cyanobacteria*

Anoxygenic phototrophs

- Phylum *Chlorobi*
- Phylum *Chloroflexi*; class *Chloroflexi*
- Phylum *Proteobacteria*; class *Gammaproteobacteria*; order *Chromatiales*
- Phylum *Proteobacteria*; class *Alphaproteobacteria*; order *Rhodospirillales*; family *Rhodospirillaceae*

Degraders

- Phylum *Bacteroidetes*
- Phylum *Firmicutes*
- Phylum *Planctomycetes*
- Phylum *Spirochaetes*

- Phylum *Fibrobacteres*

Sulphate reducers

- Phylum *Proteobacteria*; class *Deltaproteobacteria*; order *Desulfovibrionales*
- Phylum *Proteobacteria*; class *Deltaproteobacteria*; order *Desulfobacterales*
- Phylum *Proteobacteria*; class *Deltaproteobacteria*; order *Desulfuromonadales*
- Phylum *Proteobacteria*; class *Deltaproteobacteria*; order *Desulfarculales*
- Phylum *Proteobacteria*; class *Deltaproteobacteria*; order *Thermodesulfobacteriales*
- Phylum *Nitrospirae*; class *Nitrospira*; order *Nitrospirales*; family *[Thermodesulfovibrionaceae]*
- Phylum *Firmicutes*; class *Clostridia*; order *Clostridiales*; family *Peptococcaceae*; genus *Desulfotomaculum*
- Phylum *Proteobacteria*; class *Deltaproteobacteria*; order *Syntrophobacterales*; family *Syntrophaceae*; genus *Desulfobacca*
- Phylum *Proteobacteria*; class *Deltaproteobacteria*; order *Syntrophobacterales*; family *Syntrophaceae*; genus *Desulfomonile*
- Phylum *Proteobacteria*; class *Epsilonproteobacteria*; order *Campylobacterales*; family *Campylobacteraceae*; genus *Sulfurospirillum*

Aerobic sulfide oxidisers

- Phylum *Proteobacteria*; class *Betaproteobacteria*; order *Hydrogenophilales*; family *Hydrogenophilaceae*; genus *Thiobacillus*
- Phylum *Proteobacteria*; class *Betaproteobacteria*; order *Burkholderiales*; family *Alcaligenaceae*
- Phylum *Proteobacteria*; class *Epsilonproteobacteria*; order *Campylobacterales*; family *Helicobacteraceae*; genus *Sulfurimonas*
- Phylum *Proteobacteria*; class *Epsilonproteobacteria*; order *Campylobacterales*; family *Helicobacteraceae*; genus *Sulfuricurvum*

Bibliography

[1] Food web illustration. http://www.ecologyedu.com/education_resources/what_is_a_food_chain_what_is_a_food_web.html, . Accessed: 2016-03-08.

[2] Illustration of hypervolume. http://www.zo.utexas.edu/courses/bio301/chapters/Chapter13/Chapter13.html, . Accessed: 2016-03-08.

[3] Plankton illustration. http://darwinproject.mit.edu/category/research/diversity-and-biogeography-research/, . Accessed: 2016-03-08.

[4] W. G. Abrahamson. Plant-animal interactions. McGraw Hill Inc., 1989.

[5] Martin Ackermann. A functional perspective on phenotypic heterogeneity in microorganisms. Nature Reviews Microbiology, 13(8):497–508, 2015.

[6] John F Addicott, John M Aho, Michael F Antolin, Dianna K Padilla, S John, Daniel A Soluk, and John S Richardson. Ecological neighborhoods: scaling environmental pattern. Oikos, 49:340–346, 1987.

[7] Sameer Agarwal, Josh Wills, Lawrence Cayton, Gert RG Lanckriet, David J Kriegman, and Serge Belongie. Generalized non-metric multidimensional scaling. AISTATS, pages 11–18, 2007.

[8] Ramiro Aguilar, Lorena Ashworth, Leonardo Galetto, and Marcelo Adrian Aizen. Plant reproductive susceptibility to habitat fragmentation: review and synthesis through a meta-analysis. Ecology Letters, 9(8):968–980, aug 2006.

[9] Priyanga Amarasekare. Competitive coexistence in spatially structured environments: a synthesis. Ecology Letters, 6(12):1109–1122, oct 2003. ISSN 1461023X. doi: 10.1046/j.1461-0248.2003.00530.x.

[10] Delia Castro Anderson and Rosalina V Hairston. The Winogradsky Column and Biofilms: Models for Teaching Nutrient Cycling and Succession in an Ecosystem. The American Biology Teacher, 61(6):453–459, 1999.

[11] Robert A. Armstrong and Richard McGehee. Coexistence of species competing for shared resources. Theoretical Population Biology, 9(3): 317–328, jun 1976. doi: 10.1016/0040-5809(76)90051-4. URL `http://dx.doi.org/10.1016/0040-5809(76)90051-4`.

[12] Robert A Armstrong and Richard McGehee. Competitive Exclusion. The American Naturalist, 115(2):151–170, 1980.

[13] F. J. Aznar, J. A. Balbuena, and A. Raga, J. Are epizoites biological indicators of a western Mediterranean striped dolphin die-off? . Diseases of Aquatic Organisms, 18:159–163, 1994.

[14] Jack Baker, Kris French, and Robert J. Whelan. The edge effect and ecotonal species: bird communities across a natural edge in southeastern Australia. Ecology, 83(11):3048–3059, 2002.

[15] Avi Bar-Massada, Rafi Kent, and Yohay Carmel. Environmental heterogeneity affects the location of modelled communities along the niche-neutrality continuum. Proceedings. Biological sciences / The Royal Society, 281:20133249, may 2014.

[16] Albert Barberán, Scott T Bates, Emilio O Casamayor, and Noah Fierer. Using network analysis to explore co-occurrence patterns in soil microbial communities. The ISME Journal, 6:343–351, 2011.

[17] Erez David Barkae, Inon Scharf, Zvika Abramsky, and Ofer Ovadia. Jack of all trades, master of all: a positive association between habitat niche breadth and foraging performance in pit-building antlion larvae. PloS one, 7(3):e33506, jan 2012. ISSN 1932-6203. doi: 10.1371/journal.pone.0033506.

[18] G. Bell. Neutral macroecology. Science, 293(5539):2413–2418, 2001.

[19] Edward A. Bender. Perturbation Experiments in Community Ecology: Theory and Practice. Ecology, 65(1):1–13, 1984.

[20] William M. Block, Leonard A. Brennan, and R. J. Gutiérrez. Evaluation of guild-indicator species for use in resource management. Environmental Management, 11(2):265–269, 1987.

[21] Dries Bonte, Nele Vandenbroecke, Luc Lens, and Jean-Pierre Maelfait. Low propensity for aerial dispersal in specialist spiders from fragmented landscapes. Proceedings: Biological Sciences, 270(1524):1601–1607, 2003.

[22] Laurence A Boorman. Salt marshes present functioning and future change. Mangroves and Salt Marshes, 3(4):227–241, 1999.

[23] Isabelle Boulangeat, Sébastien Lavergne, Jérémie Van Es, Luc Garraud, and Wilfried Thuiller. Niche breadth, rarity and ecological characteristics within a regional flora spanning large environmental gradients. Journal of Biogeography, 39(1):204–214, jan 2012. ISSN 03050270. doi: 10.1111/j.1365-2699.2011.02581.x.

[24] Diana E. Bowler and Tim G. Benton. Causes and consequences of animal dispersal strategies: relating individual behaviour to spatial dynamics. Biological Reviews, 80(2):205–225, 1999.

[25] Rob W. Brooker, Justin M.J. Travis, Ewen J. Clark, and Calvin Dytham. Modelling species' range shifts in a changing climate: The impacts of biotic interactions, dispersal distance and the rate of climate change. Journal of Theoretical Biology, 245(1):59–65, 2007. doi: 10.1016/j.jtbi.2006.09.033.

[26] Joel S. Brown and Noel B. Pavlovic. Evolution in heterogeneous environments: Effects of migration on habitat specialization. Evolutionary Ecology, 6(5):360–382, sep 1992.

[27] Redouan Bshary and Alexandra S. Grutter. Image scoring and cooperation in a cleaner fish mutualism. Nature, 441:975–978, 2006.

[28] Lucie Büchi and Séverine Vuilleumier. Coexistence of specialist and generalist species is shaped by dispersal and environmental factors. The American naturalist, 183(5):612–624, may 2014.

[29] Timothy Bush. Response to Environmental Perturbations in Microbial Nutrient-Cycling PhD Thesis, University of Edinburgh, 2015.

[30] Ragan M. Callaway and Steven C. Pennings. Facilitation may buffer competitive effects: Indirect and diffuse interactions among salt marsh plants. The American Naturalist, 156(4):416–424, oct 2000.

[31] J Gregory Caporaso, Justin Kuczynski, Jesse Stombaugh, Kyle Bittinger, Frederic D Bushman, Elizabeth K Costello, Noah Fierer, Antonio Gonzalez

Peña, Julia K Goodrich, Jeffrey I Gordon, Gavin A Huttley, Scott T Kelley, Dan Knights, Jeremy E Koenig, Ruth E Ley, Catherine A Lozupone, Daniel McDonald, Brian D Muegge, Meg Pirrung, Jens Reeder, Joel R Sevinsky, Peter J Turnbaugh, William A Walters, Jeremy Widmann, Tanya Yatsunenko, Jesse Zaneveld, and Rob Knight. QIIME allows analysis of high-throughput community sequencing data. Nature Methods, 7(5):335–336, apr 2010. doi: 10.1038/nmeth.f.303. URL http://dx.doi.org/10.1038/nmeth.f.303.

[32] Peter Chesson. Mechanisms of maintenance of species diversity. Annual review of Ecology and Systematics, 2000.

[33] Joanne Clavel, Romain Julliard, and Vincent Devictor. Worldwide decline of specialist species: toward a global functional homogenization? Frontiers in Ecology and the Environment, 9(4):222–228, may 2011.

[34] Miguel Clavero and Lluís Brotons. Functional homogenization of bird communities along habitat gradients: accounting for niche multidimensionality. Global Ecology and Biogeography, pages 684–696, may 2010. ISSN 1466822X. doi: 10.1111/j.1466-8238.2010.00544.x.

[35] Yehuda Cohen, Bo Barker Jorgensen, Niels Peter Revsbech, and Ricardo Poplawski. Adaptation to Hydrogen Sulfide of Oxygenic and Anoxygenic Photosynthesis among Cyanobacteria. Applied and Environmental Microbiology, 51(2):398–407, 1986.

[36] Robert K Colwell and Eduardo R Fuentes. Experimental studies of the niche. Annual review of Ecology and Systematics, 6(1975):281–310, 1975.

[37] Joseph H. Connell. Diversity in Tropical Rain Forests and Coral Reefs. Science, 199(4335):1302–1310, mar 1978.

[38] R. K. Cowen. Connectivity of marine populations: Open or closed? Science, 287(5454):857–859, 2000.

[39] K. Z. Coyte, J. Schluter, and K. R. Foster. The ecology of the microbiome: Networks, competition, and stability. Science, 350(6261):663–666, nov 2015. doi: 10.1126/science.aad2602. URL http://dx.doi.org/10.1126/science.aad2602.

[40] Shabir A Dar, Alfons J M Stams, and J Gijs Kuenen. Co-existence of physiologically similar sulfate-reducing bacteria in a full-scale sulfidogenic

bioreactor fed bwith a single organic electron donor. Applied Microbial Biotechnology, 75(4):1463–1472, 2007.

[41] Catherine M. Davey, Dan E. Chamberlain, Stuart E. Newson, David G. Noble, and Alison Johnston. Rise of the generalists: evidence for climate driven homogenization in avian communities. Global Ecology and Biogeography, 21(5):568–578, jun 2011.

[42] Catherine M Davey, Dan E Chamberlain, Stuart E Newson, David G Noble, and Alison Johnston. Rise of the generalists: evidence for climate driven homogenization in avian communities. Global Ecology and Biogeography, 21(5):568–578, may 2012.

[43] Troy Day. Competition and the effect of spatial resource heterogeneity on evolutionary diversification. The American Naturalist, 155(6):790–803, 2000.

[44] D. L. DeAngelis and Wolf M. Mooij. Individual-Based Modeling of Ecological and Evolutionary Processes. Annual Review of Ecology, Evolution, and Systematics, 36:147–168, 2005.

[45] D. L. DeAngelis and J. C. Waterhouse. Equilibrium and Nonequilibrium concepts in ecological models. Ecological Monographs, 57(1):1–21, 1987.

[46] Vincent Devictor, Romain Julliard, and Frédéric Jiguet. Distribution of specialist and generalist species along spatial gradients of habitat disturbance and fragmentation. Oikos, 17:507–514, feb 2008.

[47] Vincent Devictor, Joanne Clavel, Romain Julliard, Sébastien Lavergne, David Mouillot, Wilfried Thuiller, Patrick Venail, Sébastien Villéger, and Nicolas Mouquet. Defining and measuring ecological specialization. Journal of Applied Ecology, 47(1):15–25, feb 2010. ISSN 00218901. doi: 10.1111/j.1365-2664.2009.01744.x.

[48] E. S. Dierenfeld, H. F. Hintz, J. B. Robertson, P. J Van Soest, and O. T. Oftedal. Utilization of Bamboo by the Giant Panda. The Journal of Nutrition, 112:636–641, 1982.

[49] Daniel Dykhuizen and Maxine Davies. An Experimental Model : Bacterial Specialists and Generalists Competing in Chemostats. Ecology, 61(5):1213–1227, 1980.

[50] Martin H. Entling, Karin Stämpfli, and Otso Ovaskainen. Increased propensity for aerial dispersal in disturbed habitats due to intraspecific variation and species turnover. Oikos, 120(7):1099–1109, 2010.

[51] Matthew R. Evans, Volker Grimm, Karin Johst, Tarja Knuuttila, Rogier de Langhe, Catherine M. Lessells, Martina Merz, Maureen A. O'Malley, Steve H. Orzack, Michael Weisberg, Darren J. Wilkinson, Olaf Wolkenhauer, and Tim G. Benton. Do simple models lead to generality in ecology? Trends in Ecology & Evolution, 28(10):578–583, oct 2013.

[52] Lenore Fahrig, Jacques Baudry, Lluís Brotons, Françoise G Burel, Thomas O Crist, Robert J Fuller, Clelia Sirami, Gavin M Siriwardena, and Jean-Louis Martin. Functional landscape heterogeneity and animal biodiversity in agricultural landscapes. Ecology letters, 14(2):101–112, feb 2011. ISSN 1461-0248.

[53] Meng Fan and Ke Wang. Global Periodic Solutions of a Generalized n-Species Gilpin-Ayala Competition Model . Computers and Mathematics with Applications, 40:1141–1151, 2000.

[54] D. O. Fisher, S. P. Blomberg, and I. P. F. Owens. Extrinsic versus intrinsic factors in the decline and extinction of australian marsupials. Proceedings of the Royal Society B: Biological Sciences, 270(1526):1801–1808, 2003.

[55] Richard T T Forman. Some general principles of landscape and regional ecology. Landscape Ecology, 10(3):133–142, 1995.

[56] Rima B Franklin, Jay L Garland, Carl H Bolster, and Aaron L Mills. Impact of Dilution on Microbial Community Structure and Functional Potential: Comparison of Numerical Simulations and Batch Culture Experiments. Applied and Environmental Microbiology, 67(2):702–712, 2001.

[57] Guillaume Fried, Sandrine Petit, and Xavier Reboud. A specialist-generalist classification of the arable flora and its response to changes in agricultural practices. BMC Ecology, 10(1):20, 2010.

[58] James D. Fry. Detecting ecological trade-offs using selection experiments. Ecology, 84(7):1672–1678, 2003.

[59] Douglas J Futuyma and Gabriel Moreno. The evolution of ecological specialization. Annual Review of Ecology and Systematics, 19(20):207–233, 1988.

[60] Carol X. Garzon-Lopez, Liliana Ballesteros-Mejia, Alejandro Ordoñez, Stephanie A. Bohlman, Han Olff, and Patrick A. Jansen. Indirect interactions among tropical tree species through shared rodent seed predators: a novel mechanism of tree species coexistence. Ecology Letters, 18(8):752–760, 2015.

[61] John Gillespie. Polymorphism in Patchy Environments. The American naturalist, 108(960):145–151, 1974.

[62] Josè M. Gòmez and Adela Gonzàlez-Megìas. Asymmetrical interactions between ungulates and phytophagous insects: being different matters. Ecology, 83, 2002.

[63] James R. Gosz and Peter J. H. Sharpe. Broad-scale concepts for interactions of climate, topography, and biota at biome transitions. Landscape Ecology, 3(3-4):229–243, dec 1989. ISSN 0921-2973. doi: 10.1007/BF00131541.

[64] Dominique Gravel, Charles D Canham, Marilou Beaudet, and Christian Messier. Reconciling niche and neutrality: the continuum hypothesis. Ecology letters, 9(4):399–409, apr 2006.

[65] Garrett Hardin. The competitive exclusion principle. Science, 3409:1292–1297, 1960.

[66] Caroline S. Harwood. Ecology of Spirochetes. Annual Review of Microbiology, 38:161–192, 1984.

[67] Doeke R. Hekstra and Stanislas Leibler. Contingency and statistical laws in replicate microbial closed ecosystems. Cell, 149(5):1164–1173, may 2012. doi: 10.1016/j.cell.2012.03.040. URL `http://dx.doi.org/10.1016/j.cell.2012.03.040`.

[68] Janne Heliölä, Matti Koivula, and Jari Niemelä. Distribution of Carabid Beetles (Coleoptera , Carabidae) across a Boreal Forest Clearcut Ecotone. Conservation Biology, 15(2):370–377, 2001.

[69] Holmes1961. Effects of Concurrent Infections on Hymenolepis diminuta (Cestoda) and Moniliformis dubius (Acanthocephala). I. General Effects and Comparison with Crowding. Journal of Parasitology, 47(2):209–216, 1961.

[70] R HOLT. Emergent neutrality. Trends in Ecology & Evolution, 21(10): 531–533, 2006.

[71] S P Hubbell. The Unified Neutral Theory of Biodiversity and Biogeography. Princeton University Press, USA, 2001.

[72] Edward M Hulburt. Adaptation and niche breadth of phytoplankton species along a nutrient gradient in the ocean. Journal of Plankton Research, 7(4): 581–594, 1985.

[73] G. E. Hutchinson. Homage to Santa Rosalia or Why are there so many kinds of animals? The American Naturalist, 93(870):145–159, 1959.

[74] G. E. Hutchinson. The paradox of the plankton. The American Naturalist, 95(882):137–145, 1961.

[75] John L. Ingraham, Ole Maale, and Frederick C. Neidhardt. Growth of the Bacterial Cell. Sinauer Associates Inc, 1983.

[76] Marianne Iversen, Jon Aars, Tore Haug, Inger G. Alsos, Christian Lydersen, Lutz Bachmann, and Kit M. Kovacs. The diet of polar bears (Ursus maritimus) from Svalbard, Norway, inferred from scat analysis. Polar Biology, 36(4):561–571, 2013.

[77] Erik Jeppesen, Martin Søndergaard, Brian Kronvang, Jens P Jensen, Lars M Svendsen, and Torben L Lauridsen. Lake and catchment management in Denmark. Hydrobiologia, 395:419–432, 1999.

[78] Christine M Jessup, Rees Kassen, Samantha E. Forde, Ben Kerr, Angus Buckling, Paul B. Rainey, and Brendan J.M. Bohannan. Big questions, small worlds: microbial model systems in ecology. Trends in Ecology and Evolution, 19(4):189–197, 2004.

[79] C G Jones, J H Lawton, and M Shachak. Organisms as ecosystem engineers. Oikos, 69:373–386, 1994.

[80] Clive G Jones, John H Lawton, and Moshe Shachak. Positive and Negative effects of Organisms as Physical Ecosystem Engineers. Ecology, 78(7):1946–1957, 1997.

[81] Ian D Jonsen and Lenore Fahrig. Response of generalist and specialist insect herbivores to landscape spatial structure. Lanscape Ecology, 12(902):185–197, 1997.

[82] Romain Julliard, Joanne Clavel, Vincent Devictor, Frédéric Jiguet, and Denis Couvet. Spatial segregation of specialists and generalists in bird communities. Ecology Letters, 9(11):1237–1244, nov 2006.

[83] Ronen Kadmon and Omri Allouche. Integrating the effects of area, isolation, and habitat heterogeneity on species diversity: a unification of island biogeography and niche theory. The American naturalist, 170(3):443–54, sep 2007. ISSN 1537-5323.

[84] R Kassen. The experimental evolution of specialists, generalists, and the maintenance of diversity. Journal of Evolutionary Biology, 15(2):173–190, 2002.

[85] Michael H. Kido. Food relations between coexisting native Hawaiian stream fishes. Environmental Biology of Fishes, 49(4):481–494, 1997.

[86] J. F. C. Kingman. On the genealogy of large populations. Journal of Applied Probability, 19:27, 1982.

[87] H. Kudo, K. J. Cheng, and J. W. Costerton. Interactions between Treponema bryantii and cellulolytic bacteria in the in vitro degradation of straw cellulose. Canadian Journal of Microbiology, 33:244–248, 1987.

[88] Lauri Laanisto, Riin Tamme, Inga Hiiesalu, Robert Szava-Kovats, Antonio Gazol, and Meelis Pärtel. Microfragmentation concept explains non-positive environmental heterogeneity-diversity relationships. Oecologia, 171(1):217–26, jan 2012.

[89] Peter B. Landres, Jared Verner, and Jack Ward Thomas. Ecological Uses of Vertebrate Indicator Species: A Critique. Conservation Biology, 2(4): 316–328, 1988.

[90] Mathew A. Leibold and Mark A. McPeek. Coexistence of the niche and neutral perspectives in community ecology. Ecology, 87(6):1399–1410, 2006.

[91] Jonathan M. Levine. Indirect facilitation: evidence and predictions from a riparian community. Ecology, 80(5):1762–1769, jul 1999.

[92] Richard Levins. Theory of Fitness in a Heterogeneous Environment . I . The Fitness Set and Adaptive Function. The American Naturalist, 96(891): 361–373, 1962.

[93] Richard Levins. Evolution in changing environments: some theoretical explorations. Princeton University Press, 1968.

[94] Jeremy T Lundholm and Douglas W Larson. Relationships between spatial environmental heterogeneity and plant species diversity on a limestone pavement. Ecography, 6(April):715–722, 2003.

[95] Michael Lynch and Wilfried Gabriel. Environmental Tolerance. The American naturalist, 129(2):283–303, 1986.

[96] Robert MacArthur and Edward O. Wilson. The Theory of Island Biogeography. Princeton University Press, 1967.

[97] Robert H. MacArthur and Eric R. Pianka. On Optimal Use of a Patchy Environment. The American Naturalist, 100(916):603–609, 1966.

[98] David W. Macdonald. The ecology of carnivore social behaviour. Nature, 301(3):379–383, 1983.

[99] R Mahadevan and D R Lovley. The Degree of Redundancy in Metabolic Genes Is Linked to Mode of Metabolism. Biophysical Journal, 94(4):1216–1220, 2008.

[100] Julian R. Marchesi and Jacques Ravel. The vocabulary of microbiome research: a proposal. Microbiome, 3, 2015.

[101] Larry G Marshall. Land Mammals and the Great American Interchange. American Scientist, 76(4):380–388, 1988.

[102] Robert M May and Robert H MacArthur. Niche Overlap as a Function of Environmental Variability. Proceedings of the National Academy of Sciences of the United States of America, 69(5):1109–1113, 1972.

[103] Declan J. McCabe. Sampling Biological Communities. Nature Education Knowledge, 3:63, 2011.

[104] M-l Meijer, E Jeppesen, E Van Donk, B Moss, M Scheffer, E Lammens, and E Van Nes. Long-term responses to fish-stock reduction in small shallow lakes : interpretation of five-year results of four biomanipulation cases in The Netherlands and Denmark. Hydrobiologia, pages 457–466, 1994.

[105] Ronald I. Miller. Mapping the Diversity of Nature. Springer Netherlands, 1994.

[106] Scott R Miller and Brad M Bebout. Variation in Sulfide Tolerance of Photosystem II in Phylogenetically Diverse Cyanobacteria from Sulfidic Habitats. Applied and Environmental Microbiology, 70(2):736–744, 2003.

[107] Thomas E. Miller. Direct and Indirect Species Interactions in an Early Old-Field Plant Community. The American Naturalist, 143(6):1007–1025, 1994.

[108] William J Mitsch, Xinyuan Wu, Robert W Nairn, Paul E Weihe, Naiming Wang, Robert Deal, and Charles E Boucher. Creating and Restoring Wetlands. BioScience, 48(12):1019–1027, 1998.

[109] Simon Mole and Anthony J. Zera. Differential allocation of resources underlies the dispersal-reproduction trade-off in the wing-dimorphic cricket, gryllus rubens. Oecologia, 93:121–127, 1993.

[110] Emilio Montesinos, Ricardo Guerrero, Carlos Abella, and Isabel Esteve. Ecology and Physiology of the Competition for Light Between Chlorobium limicola and Chlorobium phaeobacteroides in Natural Habitats. Applied and Environmental Microbiology, 46(5):1007–1016, 1983.

[111] Douglas W Morris. Coexistence of Specialist and Generalist Rodents Via Habitat Selection. Ecology, 77(8):2352–2364, 1996.

[112] Philip L. Munday. Habitat loss, resource specialization, and extinction on coral reefs. Global Change Biol, 10(10):1642–1647, oct 2004.

[113] James D. Murray. Mathematical Biology: I. An Introduction (Interdisciplinary Applied Mathematics) (Pt. 1). Springer, 2002.

[114] C. D. Nadell and B. L. Bassler. A fitness trade-off between local competition and dispersal in vibrio cholerae biofilms. Proceedings of the National Academy of Sciences, 108(34):14181–14185, aug 2011.

[115] Cornelis J Nagelkerke and Steph B J Menken. Coexistence of habitat specialists and generalists in metapopulation models of multiple-habitat landscapes. Acta Biotheoretica, 61:467–80, 2013.

[116] Wolfgang Nentwig. Non-webbuilding spiders: prey specialists or generalists? Oecologia, 69(4):571–576, 1986.

[117] Andrew E. Noble and William F. Fagan. A niche remedy for the dynamical problems of neutral theory. Theoretical Ecology, 8(1):149–161, nov 2014.

[118] Moriya Ohkuma, Satoko Noda, and Toshiaki Kudo. Phylogenetic Diversity of Nitrogen Fixation Genes in the Symbiotic Microbial Community in the Gut of Diverse Termites. Applied and environmental microbiology, 65(11): 4926–4934, 1999.

[119] Jari Oksanen and Peter R Minchin. Continuum theory revisited: what shape are species responses along ecological gradients? Ecological Modelling, 157(2-3):119–129, nov 2002. doi: 10.1016/s0304-3800(02)00190-4. URL http://dx.doi.org/10.1016/S0304-3800(02)00190-4.

[120] A Mark Osborn and Cindy J Smith. Molecular Microbial Ecology. Taylor and Francis Group, Abingdon, 2005.

[121] Eulyn Pagaling, Fiona Strathdee, Bryan M Spears, Michael E Cates, Rosalind J Allen, and Andrew Free. Community history affects the predictability of microbial ecosystem development. The ISME Journal, 8: 19–30, 2014.

[122] Eulyn Pagaling, Kristin Vassileva, Catherine G Mills, Timothy Bush, Richard A Blythe, Jana Schwarz-Linek, Fiona Strathdee, Rosalind J Allen, and Andrew Free. Variability in Microbial Community Structure in Replicate Nutrient-Cycling Model Ecosystems. (in preparation), 2016.

[123] Kalle Parvinen and Martijn Egas. Dispersal and the evolution of specialisation in a two-habitat type metapopulation. Theoretical Population Biology, 66(3):233–248, 2004.

[124] Bryan Patterson and Rosendo Pascual. The Fossil Mammal Fauna of South America. The Quarterly Review of Biology, 43:409–451, 1968.

[125] Michael J L Peers, Daniel H Thornton, and Dennis L Murray. Reconsidering the specialist-generalist paradigm in niche breadth dynamics: resource gradient selection by Canada lynx and bobcat. PloS one, 7(12):e51488, jan 2012. ISSN 1932-6203. doi: 10.1371/journal.pone.0051488.

[126] H Ronald Pulliam. On the relationship between niche and distribution. Ecology Letters, 3(4):349–361, jul 2000.

[127] Steven Railsback, Mark Gard, Bret C. Harvey, Jason L. White, and Julie K. H. Zimmerman. Contrast of Degraded and Restored Stream Habitat Using an Individual-Based Salmon Model. North American Journal of Fisheries Management, 33:384–399, 2013.

[128] Xavier Reboud and Graham Bell. Experimental evolution in Chlamydomonas. III. Evolution of specialist and generalist types in environments that vary in space and time. Heredity, 78(5):507–514, 1997.

[129] Mark Rees. Trade-offs among dispersal strategies in british plants. Nature, 366(6451):150–152, nov 1993.

[130] Paul G Risser. The Status of the Science Examining Ecotones between more homogeneous vegetation associations. Bioscience, 45(5):318–325, 1995.

[131] Klaus Rohde. Latitudinal gradients in species the search for the diversity : primary cause. Oikos, 65:514–527, 1992.

[132] Ophélie Ronce and Mark Kirkpatrick. When sources become sinks: Migrational meltdown in heterogeneous habitats. Evolution, 55(8):1520–1531, 2001.

[133] Philip Ronia, Timothy J Beechiea, Robert E Bilbya, Frank E Leonettia, Michael M Pollocka, and George R Pessa. A Review of Stream Restoration Techniques and a Hierarchical Strategy for Prioritizing Restoration in Pacific Northwest Watersheds. North American Journal of Fisheries Management, 22(1):1–20, 2002.

[134] Thomas P. Rooney, Shannon M. Wiegmann, David A. Rogers, and D. M. Waller. Biotic impoverishment and homogenization in unfragmented forest understory communities. Conservation Biology, 18(3):787–798, jun 2004.

[135] Madigan R.T., Martinko J., Stahl D A, and Clark D P. Brock Biology of Microorganisms, 13th edn. Pearson Prentice Hall, London, 2011.

[136] M Scheffer, S Hosper, M Meijer, B Moss, and E Jeppesen. Alternative equilibria in shalow lakes. Trends in Ecology and Evolution, 8(8):275–279, 1993.

[137] M Scheffer, S Carpenter, J a Foley, C Folke, and B Walker. Catastrophic shifts in ecosystems. Nature, 413(6856):591–6, 2001.

[138] Marten Scheffer and Egbert H. van Nes. Self-organized similarity, the evolutionary emergence of groups of similar species. Proceedings of the National Academy of Sciences of the United States of America, 103:6230–6235, 2006.

[139] David W. Schindler, R. E. Hecky, D. L. Findlay, M. P. Stainton, B. R. Parker, M. J. Paterson, K. G. Beaty, M. Lyng, and S. E. M. Kasian. Eutrophication of lakes cannot be controlled by reducing nitrogen input: Results of a 37-year whole-ecosystem experiment. Proceedings of the National Academy of Sciences of the United States of America, 105(32): 11254–11258, 2008.

[140] T W Schoener. Resource partitioning in ecological communities. Science (New York, N.Y.), 185(4145):27–39, jul 1974.

[141] J W Schopf. Disparate rates, differing fates: tempo and mode of evolution changed from the Precambrian to the Phanerozoic. Proceedings of the National Academy of Sciences of the United States of America, 91(15): 6735–6742, 1994.

[142] Dylan W. Schwilk and David D. Ackerly. Limiting similarity and functional diversity along environmental gradients. Ecology Letters, 8(3):272–281, jan 2005.

[143] Daniel Simberloff and Tamar Dayan. The Guild Concept and the Structure of Ecological Communities. Annual Review of Ecology and Systematics, 22:115–143, 1991.

[144] Francis J. Singer. Wild pig populations in the National Parks. Environmental Management, 5(3):263–270, 1981.

[145] Douglas G. Sprugel. Disturbance, Equilibrium, and Environmental Variability: What is 'Natural' Vegetation in a Changing Environment? . Biological Conservation, 58:1–18, 1991.

[146] Jostein Starrfelt and Hanna Kokko. Bet-hedging-a triple trade-off between means, variances and correlations. Biological Reviews, 87(3):742–755, 2012.

[147] Richard A. Stillman, Steven F. Railsback, Jarl Giske, Uta Berger, and Volker Grimm. Making Predictions in a Changing World: The Benefits of Individual Based Ecology. BioScience, 65(2):140–150, 2015.

[148] Didier Stilmant, Cécile Van Bellinghen, Thierry Hance, and Guy Boivin. Host specialization in habitat specialists and generalists. Oecologia, 156(4): 905–912, 2008.

[149] Erik Stockstad. Canada's Experimental Lakes. Science, 322:1316–1319, 2008.

[150] Maayke Stomp, Jef Huisman, Lucas J Stal, and Hans C P Matthijs. Colorful niches of phototrophic microorganisms shaped by vibrations of the water molecule. The ISME Journal, 1(4):271–282, aug 2007.

[151] Lewi Stone and Alan Roberts. Conditions for a species to gain advantage from the presence of competitors. Ecology, 72(6):1964, 1991.

[152] Garret Suen, Paul J Weimer, David M Stevenson, Frank O Aylward, Julie Boyum, Jan Deneke, Colleen Drinkwater, Natalia N Ivanova, Natalia Mikhailova, Olga Chertkov, Lynne A Goodwin, Cameron R Currie, David Mead, and Phillip J Brumm. The Complete Genome Sequence of Fibrobacter succinogenes S85 Reveals a Cellulolytic and Metabolic Specialist. Plos One, 6(4):e18814, 2001.

[153] T E Tew, I a Todd, and D W Macdonald. Arable habitat use by wood mice (Apodemus sylvaticus). 2. Microhabitat. Journal of Zoology, 250(3): 305–311, 2000.

[154] Ken Thompson, John G Hodgeson, and Kevin J Gaston. Abundance-range size relationships in the herbaceous flora of central England. Journal of Ecology, 86:439–448, 1998.

[155] David Tilman. The Importance of the Mechanisms of Interspecific Competition. The American Naturalist, 129(5):769–774, 1987.

[156] David Tilman. Niche tradeoffs, neutrality, and community structure: a stochastic theory of resource competition, invasion, and community assembly. Proceedings of the National Academy of Sciences of the United States of America, 101(30):10854–61, jul 2004.

[157] Justin M. J. Travis, Karen Mustin, Kamil A. Bartoń, Tim G. Benton, Jean Clobert, Maria M. Delgado, Calvin Dytham, Thomas Hovestadt, Stephen C. F. Palmer, Hans Van Dyck, and Dries Bonte. Modelling dispersal: an eco-evolutionary framework incorporating emigration, movement, settlement behaviour and the multiple costs involved. Methods in Ecology and Evolution, 3(4):628–641, feb 2012. doi: 10.1111/j.2041-210x.2012.00193.x. URL http://dx.doi.org/10.1111/j.2041-210X.2012.00193.x.

[158] Peter J. Turnbaugh, Ruth E. Ley, Michael A. Mahowald, Vincent Magrini, Elaine E. Mardis, and Jeffrey I. Gordon. An obesity-associated gut

microbiome with increased capacity for energy harvest. Nature, 444:1027–1031, 2006.

[159] Peter J. Turnbaugh, Fredrik Backhed, Lucinda Fulton, and Jeffrey I. Gordon. Diet-Induced Obesity Is Linked to Marked but Reversible Alterations in the Mouse Distal Gut Microbiome. Cell Host and Microbe, 3:213–223, 2008.

[160] Peter J. Turnbaugh, Micah Hamady, Tanya Yatsunenko, Brandi L. Cantarel, Alexis Duncan, Ruth E. Ley, Mitchell L. Sogin, William J. Jones, Bruce A. Roe, Jason P. Affourtit, Michael Egholm, Bernard Henrissat, Andrew C. Heath, Rob Knight, and Jeffrey I. Gordon. A core gut microbiome in obese and lean twins. Nature, 457:480–484, 2009.

[161] Eddy van der Maarel. Ecotones and Ecoclines are Different. Journal of Vegetation Science, 1(1):135–138, 1990.

[162] Peter H van Tienderen. Evolution of Generalists and Specialist in Spatially Heterogeneous Environments. Evolution, 45(6):1317–1331, 1991.

[163] Diego P Vázquez and Daniel Simberloff. Ecological specialization and susceptibility to disturbance: conjectures and refutations. The American naturalist, 159(6):606–23, jun 2002.

[164] I. Le Viol, F. Jiguet, L. Brotons, S. Herrando, A. Lindstrom, J. W. Pearce-Higgins, J. Reif, C. Van Turnhout, and V. Devictor. More and more generalists: two decades of changes in the european avifauna. Biology Letters, 8(5):780–782, jul 2012.

[165] Peter J Wangersky. Lotka-Volterra Population Models. Annual Review of Ecology and Systematics, 9:189–218, 1978.

[166] J Weiner. The beak of the finch. Knopf (Random House), New York, 1994.

[167] Jacob Weiner and Sa Xiao. Variation in the degree of specialization can maintain local diversity in model communities. Theoretical Ecology, 5(2): 161–166, jan 2012.

[168] John T. Wertz, Eunji Kim, John A. Breznak, Thomas M. Schmidt, and Jorge L. M. Rodrigues. Genomic and Physiological Characterization of the Verrucomicrobia Isolate Diplosphaera colitermitum gen. nov., sp. nov., Reveals Microaerophily and Nitrogen Fixation Genes. Applied and Environmental Microbiology, 78:1544–1555, 2012.

[169] Eve M. White, John C. Wilson, and Anthony R. Clarke. Biotic indirect effects: a neglected concept in invasion biology. Diversity and Distributions, 12(4):443–455, jul 2006.

[170] Henry M. Wilbur. Competition, Predation, and the Structure of the Ambystoma-Rana Sylvatica Community. Ecology, 53(1):3–21, 1972.

[171] David Sloan Wilson and Jin Yoshimura. On the Coexistance of Specialists and Generalists. The American Naturalist, 144(4):692–707, 1994.

[172] J. Bastow Wilson. Mechanisms of species coexistence: twelve explanations for Hutchinson's 'paradox of the plankton': evidence from New Zealand plant communities. New Zealand Journal of Ecology, 13(1):17–42, 1990.

[173] Shaun K Wilson, Scott C Burgess, Alistair J Cheal, Mike Emslie, Rebecca Fisher, Ian Miller, Nicholas V C Polunin, P Hugh, and A Sweatman. Habitat utilization by coral reef fish : implications for specialists vs . generalists in a changing environment. Journal of Animal Ecology, 77: 220–228, 2008.

[174] Christian Winter, Thierry Bouvier, Markus G. Weinbauer, and T. Frede Thingstad. Trade-Offs between Competition and Defense Specialists among Unicellular Planktonic Organisms: the Killing the Winner Hypothesis Revisited. Microbiology and Molecular Biology Reviews, 74(1):42–57, 2010.

[175] Justin P. Wright, Clive G. Jones, and Alexander S. Flecker. An ecosystem engineer, the beaver, increases species richness at the landscape scale. Oecologia, 132(1):96–101, 2002.

[176] Takeshi Yamada, Yuji Sekiguchi, Satoshi Hanada, Hiroyuki Imachi, Akiyoshi Ohashi, Hideki Harada, and Yoichi Kamagata. Anaerolinea thermolimosa sp. nov., Levilinea saccharolytica gen. nov., sp. nov. and Leptolinea tardivitalis gen. nov., sp. nov., novel filamentous anaerobes, and description of the new classes Anaerolineae classis nov. and Caldilineae classis nov. in the bacterial phylum Chloroflexi. International Journal of Systematic and Evolutionary Microbiology, 56:1331–1340, 2006.

[177] Yoshimi Yoshino, Tobias Galla, and Kei Tokita. Rank abundance relations in evolutionary dynamics of random replicators. Physical Review E, 78(3), sep 2008. doi: 10.1103/physreve.78.031924. URL `http://dx.doi.org/10.1103/PhysRevE.78.031924`.

[178] H. Zhao, J.-R. Yang, H. Xu, and J. Zhang. Pseudogenization of the umami taste receptor gene tas1r1 in the giant panda coincided with its dietary switch to bamboo. Molecular Biology and Evolution, 27(12):2669–2673, 2010.